

Journal of Integrated

OMICS

a methodological journal

Editors-in-Chief

Carlos Lodeiro-Espiño

Florentino Fdez-Riverola

Jens Coorssen

Jose-Luís Capelo-Martínez

JIOMICS

Journal of Integrated OMICS

Focus and Scope

Journal of Integrated OMICS, JIOMICS, provides a forum for the publication of original research papers, preliminary communications, technical notes and critical reviews in all branches of pure and applied "-omics", such as genomics, proteomics, lipidomics, metabolomics or metallomics. The manuscripts must address methodological development. Contributions are evaluated based on established guidelines, including the fundamental nature of the study, scientific novelty, and substantial improvement or advantage over existing technology or method. Original research papers on fundamental studies, and novel sensor and instrumentation development, are especially encouraged. It is expected that improvements will also be demonstrated within the context of (or with regard to) a specific biological question; ability to promote the analysis of molecular mechanisms is of particular interest. Novel or improved applications in areas such as clinical, medicinal and biological chemistry, environmental analysis, pharmacology and materials science and engineering are welcome.

Editors-in-Chief

Carlos Lodeiro-Espiño, University of Vigo, Spain

Florentino Fdez-Riverola, University of Vigo, Spain

Jens R. Coorssen, University of Western Sydney, NSW, Australia

Jose-Luís Capelo-Martínez, University of Vigo, Spain

Regional editors

ASIA

Gary Xiao

Director of Functional Genomics and Proteomics Laboratories at Osteoporosis Research Center, Creighton University Omaha, Nebraska, USA

Yogeshwer Shukla

Proteomics laboratory at Indian Institute of Toxicology Research (Council of Scientific and Industrial Research), Lucknow, India

AUSTRALIA AND NEW ZEALAND

Jens R. Coorssen

University of Western Sydney, NSW, Australia

Europe

Gilberto Igrejas

University of Trás-os-Montes and Alto Douro, Life Sciences and Environmental School, Institute for Biotechnology and Bioengineering, Centre of Genetics and Biotechnology
Department of Genetics and Biotechnology, 5001-801 Vila Real, Portugal

UFZ, Helmholtz-Centre for Environmental Research, Department of Proteomics, Permoserstr. 15, 04318 Leipzig, Germany

Jan Ottervald

Research and Development | Innovative Medicines Neuroscience, CNSP iMed Science Södertälje, AstraZeneca, Sweden

Martin von Bergen

North America

Randen Patterson

Center for Computational Proteomics, The Pennsylvania State University,

USA

School of Medicine, The University of North Carolina at Chapel Hill, USA

Oscar Alzate

Associate Professor of Cell and Developmental Biology, Adjunct Associate Professor in Neurology, Director: Systems Proteomics Center

Yue Ge

US Environmental Protection Agency, Research Triangle Park, USA

South America

Eduardo Alves de Almeida

Depto. de Química e Ciências Ambientais, IBILCE - UNESP, Brazil

University of Campinas - Unicamp

Marco Aurélio Zezzi Arruda

Carlos H. I. Ramos

Chemistry Institute – UNICAMP, Brazil

Associated editors

AFRICA

Saffaj Taouqif

Centre Universitaire Régional d'Interface, Université Sidi Mohamed Ben Abdallah, route d'Imouzzar-Fès, Morocco

ASIA

Abdul Jaleel A

Rajiv Gandhi Centre for Biotechnology, Thycaud PO, Trivandrum, Kerala, India

Ali A. Ensafi

Isfahan University of Technology, Iran

Allison Stelling

Dresden, Germany

Amita Pal

Division of Plant Biology, Bose Institute, Kolkata, India

Ashish Gupta

Centre of Biomedical Magnetic Resonance, SGPGIMS Campus, Lucknow, India

Canhua Huang

The State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, PR China

Chaminda Jayampath Seneviratne

Oral Biosciences, Faculty of Dentistry, University of Hong Kong, Hong Kong

Cheolju Lee

Korea Institute of Science and Technology, Seoul, Korea

Chi Chiu Wang

Department of Obstetrics & Gynaecology, Chinese University of Hong Kong, Hong Kong

Chii-Shiang Chen

National Museum of Marine Biology and Aquarium, Checheng, Pingtung, Taiwan

Ching-Yu Lin

Institute of Environmental Health, College of Public Health, National Taiwan University, Taipei, Taiwan

Chantragan Srisomsap

Chulabhorn Research Institute, Bangkok, Thailand

Chen Han-Min

Department of Life Science, Catholic Fu-Jen University, Taipei, Taiwan

David Yew

Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Debmalya Barh

Institute of Integrative Omics and Applied Biotechnology (IIOAB), India

Dwaipayan Bharadwaj

Genomics & Molecular Medicine Unit, Institute of Genomics & Integrative Biology (CSIR), Mall Road, Delhi, India

Eiji Kinoshita

Department of Functional Molecular Science, Graduate School of Biomedical Sciences, Hiroshima University, Japan

Eun Joo Song

Molecular Recognition Research Center, Korea Institute of Science & Technology, Seoul, Korea

Fan Chen

Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (CAS), China

Feng Ge

Institute of Hydrobiology, Chinese Academy of Sciences, China

Ganesh Chandra Sahoo

BioMedical Informatics Center of Rajendra Memorial Research Institute of Medical Science (RMRIMS), Patna, India

Guangchuang Yu

Institute of Life & Health Engineering, Jinan University, Guangzhou, China

Gufeng Wang

Department of Chemistry, North Carolina State University, Raleigh, USA

Hai-Lei Zheng

School of Life Sciences, Xiamen University, China

Hee-bal Kim

Department of Food and Animal Biotechnology of the Seoul National University, Korea

Hsin-Yi Wu

Institute of Chemistry, Academia Sinica, Taiwan

Hitoshi Iwahashi

Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Japan

Hong-Lin Chan

National Tsing-Hua University, Taiwan

Hongying Zhong

College of Chemistry, Central China Normal University, Wuhan, P. R. China

Huan-Tsung Chang

Department of Chemistry, National Taiwan University, Taipei, Taiwan

Hua Xu

Research Resources Center, University of Illinois, Chicago

Hui-Fen Wu

Department of Chemistry, National Sun Yat – Sen University, 70, Lien-Hai Road, 80424, Kaohsiung, Taiwan

Hye-Sook Kim

Faculty of Pharmaceutical Sciences, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Japan

Hyun Joo An

ChungNam National University, Daejeon, Korea (South)

Ibrokhim Abdurakhmonov

Institute of Genetics and Plant experimental Biology Academy of Sciences of Uzbekistan, Uzbekistan

Isam Khalaila

Biotechnology Engineering Department, Ben-Gurion University, Israel

Jagannadham Medicharla

Senior Principal Scientist, CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India

Jianghao Sun

Food Composition and Method Development Lab, U.S. Dept. of Agriculture, Agricultural Research Services, Beltsville, USA

Jong Won Yun

Dept. of Biotechnology, Kyungsan, Kyungbuk 712-714, Republic of Korea

Juan Emilio Palomares-Rius

Forestry and Forest Products Research Institute, Tsukuba, Japan

Jung Min Kim

Liver and Immunology Research Center, Daejeon Oriental Hospital of Daejeon University, Republic of Korea

Kazuaki Kakehi

School of Pharmacy, Kinki University, Kowakae 3-4-1, Higashi-Osaka, 577-8502, Japan

Kazuki Sasaki

Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center, Japan

Ke Lan

West China School of Pharmacy, Sichuan University, Chengdu, China

Kelvin Leung

Department of Chemistry, Hong Kong Baptist University, Hong Kong

Kobra Pourabdollah

Razi Chemistry Research Center (RCRC), Shahreza Branch, Islamic Azad University, Shahreza, Iran

Kohji Nagano

Chugai Pharmaceutical Co. Ltd., Japan

Koji Ueda

Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, Tokyo, Japan

Krishnakumar Menon

Amrita Center for Nanosciences and Molecular Medicine, Amrita Institute of Medical Sciences, Kochi, Kerala, India

Lakshman Samaranayake

Dean, And Chair of Oral Microbiology, University of Hong Kong, Hong Kong

Lal Rai

Molecular Biology Section, Centre of Advanced Study in Botany, Banaras Hindu University, Varanasi-221005, India

Lei Zhou

Singapore Eye Research Institute, Singapore

Li Jianke

Institute of Apicultural Research, Chinese Academy of Agricultural Science, Beijing, China, HKSAR, PR China

Ling Zheng

College of Life Sciences, Wuhan University, China

Luk John Moonching

National University of Singapore, Singapore

Mahdi Ghasemi-Varnamkhasti

Department of Agricultural Machinery Engineering, Faculty of Agriculture, Shahrekord University, Shahrekord, Iran

Manjunatha Kini

Department of Biological Sciences, National University of Singapore, Singapore

Masahiro Sugimoto

Graduate School of Medicine and Faculty of Medicine, Kyoto University Medical Innovation Center, Japan

Masaya Miyazaki

National Institute of Advanced Industrial Science and Technology, 807-1 Shuku, Tosu, Saga 841-0052, Japan

Ming-Fa Hsieh

Department of Biomedical Engineering, Chung Yuan Christian University, Taiwan

Mingfeng Yang

Key Laboratory of Urban Agriculture of Ministry of Agriculture P. R. China Beijing University of Agriculture, China

Mo Yang

Interdisciplinary Division of Biomedical Engineering, the Hong Kong Polytechnic University, Hong Kong, China

Mohammed Rahman

Center of Excellence for Advanced Materials Research (CEAMR), King Abdulaziz University, Jeddah, Saudi Arabia

Moganty Rajeswari

Department of Biochemistry, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India

Nam Hoon Cho

Dept. of Pathology, Yonsei University College of Medicine, Korea

Ningwei Zhao

Life Science & Clinical Medicine Dept. ; Shimadzu (China) Co., Ltd

Pei-Yuan Qian

Division of Life Science, Hong Kong University of Science and Technology, China

Peng Zhou

Center of Bioinformatics (COBI), Key Laboratory for NeuroInformation of Ministry of Education (KLNME), University of Electronic Science and Technology of China (UESTC)

Poh-Kuan CHONG (Shirly)

National University of Singapore, Singapore

Qian Shi

Institutes of Biomedical Sciences, Fudan University, Shanghai, China

Qionglin Liang

Tsinghua University, Beijing, China

Rakesh Mishra

Centre for Cellular and Molecular Biology, Hyderabad, India

Roger Beuerman

Singapore Eye Research Institute, Singapore

Sameh Magdeldin Mohamed

Niigata prefecture, Nishi-ku, Terao, Niigata, Japan

Sanjay Gupta

Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai, India

Sanjeeva Srivastava

Indian Institute of Technology (IIT) Bombay, India

Seiichi Uno

Education and Research Center for Marine Resources and Environment, Faculty of Fisheries, Kagoshima University, Japan

Sen-Lin Tang

Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

Setsuko Komatsu

National Institute of Crop Science, Japan

Shaojun Dai

Alkali Soil Natural Environmental Science Center, Key Laboratory of Saline-alkali Vegetation Ecology Restoration in Oil Field, Ministry of Education, Northeast Forestry University, P.R. China

Shipin Tian

Institute of Botany, Chinese Academy of Sciences, China

Songping Liang

Hunan Normal University, Changsha City, China

Steven Shaw

Department of Obstetrics and Gynecology, Chang Gung Memorial Hospital, Linkou, Taiwan

Suresh Kumar

Department of Applied Chemistry, S. V. National Institute of Technology, Gujarat, India

Tadashi Kondo

National Cancer Center Research Institute, Japan

Taesusung Park

National Research Laboratory of Bioinformatics and Biostatistics at the Department of Statistics Seoul National University, Korea

Toshihide Nishimura

Department of Surgery I, Tokyo Medical University, Tokyo, Japan

Vishvanath Tiwari

Department of Biochemistry, Central University of Rajasthan, India

Wei Wang

School of Medical Sciences, Edith Cowan University, Perth, Australia

Weichuan Yu

Department of Electronic and Computer Engineering and Division of Biomedical Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Wei-dong Zhang

Lab of Natural Products, School of Pharmacy, Second Military Medical University, Shanghai, China

Wenxiang Lin

School of Life Sciences, Fujian Agriculture and Forestry University, China

William Chen Wei Ning

School of Chemical and Biomolecular Engineering Nanyang Technological University, Singapore

Xiao LiWang

Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN

Xiao Zhiqiang

Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, P.R. China

Xiaoping Wang

Key Laboratory of Molecular Biology & Pathology, State Bureau of Chinese Medicine, China

Xuanxian Peng

School of Life Sciences, Sun Yat-sen University, Guangzhou, China

Yang Liu

Department of Chemistry, Tsinghua University, Beijing, China

YasminAhmad

Peptide and Proteomics Division Defence Institute of Physiological and Allied Research (DIPAS), DRDO, Ministry of Defence, Timarpur, Delhi-54, India

Yin Li

Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

Yong Song Gho

Department of Life Science, POSTECH, Pohang, Korea

Yoon-E Choi

Chonbuk National University, Iksan-si, South Korea

Yoon-Pin Lim

Department of Biochemistry, National University of Singapore, Singapore

Young-Gyu Ko

College of Life Sciences and Biotechnology, Korea University, Korea

Young-Suk Kim

Department of Food Science and Engineering, College of Engineering, Ewha Womans University, Seoul, Korea

Youngsoo Kim

Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea

Youxiong Que

National Research & Development Center for Sugarcane, China Agriculture Research System(CARS), Fujian Agriculture & Forestry University, Republic of China

Yu-Chang Tyan

Department of Medical Imaging and Radiological Sciences, Kaohsiung Medical University, Kaohsiung, Taiwan

Yu Wang

Department of Pharmacology and Pharmacy, the University of Hong Kong, China

Yu Xue

Department of Systems Biology, College of Life Science and Technology Huazhong University of Science and Technology, Wuhan, China

Yulan Wang

State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Centre for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, The Chinese Academy of Sciences, China

Zhengwei Yuan

The key laboratory of health ministry for congenital malformation, Shengjing Hospital, China Medical University

Zhiqiang Gao

Department of Chemistry, National University of Singapore

AUSTRALIA AND NEW ZEALAND

Bruno Catimel

Epithelial laboratory, Ludwig Institute for Cancer Research, Post Office Royal Melbourne Hospital, Australia

Daniel Cozzolino

Barley Research Laboratory, School of Agriculture, Food and Wine, University of Adelaide, Australia

David Beale

CSIRO Land and Water, Highett, Australia

Emad Kiriakous

Queensland University of Technology (QUT), Brisbane, Australia

Joëlle Coumans-Moens

School of Science and Technology, School of Medicine, University of New

England, Australia

Marc Wilkins

University of New South Wales, Sydney, Australia

Maurizio Ronci

Mawson Institute, University of South Australia, Mawson Lakes, Australia

Michelle Hill

University of Queensland, Australia

Michelle Colgrave

CSIRO Livestock Industries, St Lucia, Australia

Nicolas Taylor

ARC Centre of Excellence in Plant Energy Biology & Centre for Comparative Analysis of Biomolecular Networks (CABiN), University of Western Australia,

Perth, Australia

Peter Hoffmann

Institute for Photonics & Advanced Sensing (IPAS), School of Chemistry and Physics, University of Adelaide, Australia

Stefan Clerens

Protein Quality & Function, AgResearch Ltd Christchurch, New Zealand

Peter Solomon

Research School of Biology College of Medicine, Biology and Environment, Australian National University, Australia

Phoebe Chen

Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia

Richard Christopherson

School of Molecular Bioscience, University of Sydney, Australia

Sham Nair

Department of Biological Sciences, Faculty of Science, Macquarie University, NSW, Australia

Sylvia Urban

School of Applied Sciences (Discipline of Applied Chemistry), RMIT University, Melbourne, Victoria, Australia

Valerie Wasinger

Bioanalytical Mass Spectrometry Facility, Mark Wainwright Analytical Centre, University of NSW, Australia

Wujun Ma

Centre for Comparative Genomics, Murdoch University, Australia

Yin Xiao

Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia

EUROPE

AhmetKoc, PhD

Izmir Institute of Technology, Department of Molecular Biology & Genetics, Urla, Izmir, Turkey

Alejandro Gella

Department of Basic Sciences, Neuroscience Laboratory, Faculty of Medicine and Health Sciences, Universitat Internacional de Catalunya, Sant Cugat del Vallès-08195, Barcelona, Spain

Alessandro Pessione

Università degli Studi di Torino, Italy

Alexander Scherl

Proteomics Core Facility, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Alfio Ferlito

ENT Clinic, University of Udine, Italy

Almudena Fernández Briera

Dpt. Biochemistry Genetics and Immunology, Faculty of Biology –University of Vigo, Spain

Alfonsina D'Amato

Politecnico di Milano, Department of Chemistry, Materials and Chemical Engineering "GiulioNatta", Italy

Alfred Vertegaal

Molecular Cell Biology, Leiden University Medical Center, The Netherlands

Ali Mobasher

School of Veterinary Medicine and Science, Faculty of Medicine and Health Sciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, Leicestershire, United Kingdom

Andre Almeida

Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Portugal

Andrea Matros

Leibniz Institute of Plant Genetics and Crop Plant Research (IPK-Gatersleben), Gatersleben, Germany

Andrei Turtoi

University of Liege, Metastasis Research Laboratory, GIGA-Cancer Bât. B23, Belgium

Angelo D'Alessandro

Università degli Studi della Tuscia, Department of Ecological and Biological Sciences, Viterbo, Italy

Angelo Izzo

Department of Experimental Pharmacology, University of Naples Federico II, Naples, Italy

Antonio Gnoni

Department of Medical Basic Sciences, University of Bari "Aldo Moro", Bari, Italy

Ana Maria Rodríguez-Piñero

Institute of Biomedicine, University of Gothenburg, Sweden

Ana Varela Coelho

Instituto de Tecnologia Química e Biológica (ITQB) Universidade Nova de Lisboa (UNL), Portugal

Anna Maria Timperio

Dipartimento Scienze Ambientali Università della Tuscia Viterbo, Italy

André Nogueira Da Costa

Molecular Carcinogenesis Group, Section of Mechanisms of Carcinogenesis International Agency for Research on Cancer - World Health Organization (IARC-WHO), Lyon, France

Andreas Boehm

Steigerfurtweg 8a, D-97084 Würzburg, Germany

Andrea Scaloni

Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, via Argine 1085, 80147 Napoli, Italy

Andreas Tholey

Division for Systematic Proteome Research, Institute for Experimental Medicine, Christian-Albrechts-University, Germany

Angel Manteca

Departamento de Biología Funcional and IUBA, Facultad de Medicina, Universidad de Oviedo, Spain

Angel P. Diz

Department of Biochemistry, Genetics and Immunology, Faculty of Biology, University of Vigo, Spain

Angela Bachi

Mass Spectrometry Unit DIBIT, San Raffaele Scientific Institute, Milano, Italy

Angela Chambery

Department of Life Science, Second University of Naples, Italy

Anna-Irini Koukkou

University of Ioannina, Department of Chemistry, Biochemistry Laboratory, Greece

António Sebastião Rodrigues

Departamento de Genética, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Portugal

Arkadiusz Kosmala

Laboratory of Cytogenetics and Molecular Biology, Institute of Plant Genetics, Polish Academy of Sciences, Poland

Arzu Umar

Department of Medical Oncology, Laboratory of Breast Cancer Genomics and Proteomics, Erasmus Medical Center Rotterdam Josephine Nefkens Institute, Rotterdam, The Netherlands

Baggerman Geert

ProMeta, Interfaculty Center for Proteomics and Metabolomics, Leuven, Belgium

Bart De Spiegeleer

Ghent University, Belgium

Bart Devreese

Laboratory for Protein Biochemistry and Biomolecular Engineering,
Department for Biochemistry and Microbiology, Ghent University, Belgium

Bernard Corfe

Department of Oncology, University of Sheffield, Royal Hallamshire Hospital,
United Kingdom

Bernd Thiede

Biotechnology Centre of Oslo, University of Oslo, Blindern, Norway

Björn Meyer

Institut für Instrumentelle Analytik und Bioanalytik Hochschule Mannheim,
Germany

Bruno Baudin

Biochemistry Laboratory A, Saint-Antoine Hospital, Hôpitaux Universitaires
Est Parisien-APHP, Paris, France

Bruno Manadas

Center for Neuroscience and Cell Biology, University of Coimbra, Portugal

Cândido Pinto Ricardo

Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa,
Av. da República-EAN, 2780-157 Oeiras, Portugal

Carla Pinheiro

Plant Sciences Division, Instituto de Tecnologia Química e Biológica (ITQB),
Universidade Nova de Lisboa, Portugal

Claudia Desiderio

Consiglio Nazionale delle Ricerche, Istituto di Chimica del Riconoscimento
Molecolare (UOS Roma), Italy

Claudio De Pasquale

SAGa Department, University of Palermo, Italy

Carlos Gutiérrez Merino

Dept. Biochemistry and Molecular Biology University of Extremadura,
Badajoz, Spain

Cecilia Calado

Engineering Faculty Catholic University of Portugal, Rio de Mouro, Portugal

Celso Reis

Institute of Molecular Pathology and Immunology of the University of Porto,
IPATIMUP, Portugal

Celso Vladimiro Cunha

Medical Microbiology Department, Institute of Hygiene and Tropical
Medicine, New University of Lisbon, Portugal

Charles Steward

The Wellcome Trust Sanger Institute, Hinxton, United Kingdom

Chris Goldring

Department of Pharmacology and Therapeutics, MRC Centre for Drug Safety
Science, University of Liverpool, United Kingdom

Christian Lindermayr

Institute of Biochemical Plant Pathology, Helmholtz Zentrum München,
German Research Center for Environmental Health, Neuherberg, Germany

Christiane Fæste

Section for Chemistry and Toxicology Norwegian Veterinary Institute, Oslo,
Norway

Christer Wingren

Department of Immunotechnology, Lund University, Lund, Sweden

Christophe Cordella

UMR1145 INRA, Laboratoire de Chimie Analytique, Paris, France

Christophe Masselon

Laboratoire de Biologie a Grande Echelle (iRTSV/BGE), CEA Grenoble,
France

Cosima Damiana Calvano

Universita' degli Studi di Bari, Dipartimento di Chimica, Bari, Italy

David Cairns

Section of Oncology and Clinical Research, Leeds Institute of Molecular
Medicine, Leeds, UK

Daniela Cecconi

Dip. diBiotecnologie, Laboratori di Proteomica e Spettrometri di Massa,
Università di Verona, Verona, Italy

David Honys

Laboratory of Pollen Biology, Institute of Experimental Botany ASCR, Czech
Republic

David Sheehan

Dept. Biochemistry, University College Cork (UCC), Ireland

Deborah Penque

Departamento de Genética, Instituto Nacional de Saúde Dr Ricardo Jorge
(INSA, I.P.), Lisboa, Portugal

Dilek Battal

Mersin University, Faculty of Pharmacy, Department of Toxicology, Turkey

Domenico Garozzo

CNR ICTP, Catania, Italy

Ed Dudley

Institute of Mass Spectrometry, College of Medicine Swansea University,
Singleton Park, Swansea, Wales, UK

Edoardo Saccenti

University of Amsterdam, Netherlands Metabolomics Centre, The
Netherlands

Elena Gonzalez

Complutense University of Madrid, Dept. Biochemistry and Molecular
Biology IV, Veterinary Faculty, Madrid, Spain

Elia Ranzato

Dipartimento di Scienze e Innovazione Tecnologica, DiSIT, University of
Piemonte Orientale, Alessandria, Italy

Elisa Bona

Università del Piemonte Orientale, DISIT, Alessandria, Italy

Elke Hammer

Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-
Arndt Universität, Germany

Enrica Pessione

University of Torino, Life Sciences and Systems Biology Department, Torino,
Italy

Eva Rodríguez Suárez

Proteomics Core Facility - CIC bioGUNE, Parque tecnologico de Bizkaia,
Spain

Federica Pellati

Department of Life Sciences, University of Modena and Reggio Emilia, Italy

Ferdinando Cerciello

Laboratory of Molecular Oncology, Clinic of Oncology, University Hospital
Zürich, Switzerland

Fernando J. Corrales

Division of Hepatology and Gene Therapy, Proteomics Unit, Center for
Applied Medical Research (CIMA), Pamplona, Spain

Florian Szabados

Dept. of Medical Microbiology, Ruhr-University Bochum, Germany

Francesco Salvi

University of Milano Bicocca, Italy

Francisco J Blanco

Platform of Proteomics, Proteo-Red-ISCIII INIBIC-Hospital Universitario A
Coruña, Spain

Francisco Javier Fernández Acero

Laboratory of Microbiology, Marine and Environmental Sciences Faculty,
University of Cádiz, Pol. Río San Pedro s/n, Puerto Real, Cádiz, Spain

Francisco Torrens

Institut Universitari de Ciència Molecular, Universitat de València, Spain

François Fenaille

CEA, IBiTecS, Service de Pharmacologie et DImmunoanalyse (SPI), France

Frederic Silvestre

University of Namur, Belgium

Fulvio Magni

Department of Health Science, Monza, Italy

Georgios Theodoridis

Department of Chemistry, Aristotle University, Greece

Germain Rousselet

Laboratoire Réparation et Transcription dans les cellules Souches (LRTS), CEA/DSV/IRCM, Fontenay aux Roses, France

German Bou

Servicio de Microbiología-INIBIC, Complejo Hospitalario Universitario La Coruña, Spain

Gianfranco Mamone

Proteomic and Biomolecular Mass Spectrometry Centre, Institute of Food Science CNR, Italy

Gianfranco Romanazzi

Department of Environmental and Crop Sciences, Marche Polytechnic University, Italy

Gianluigi Mauriello

Department of Food Science, University of Naples Federico II Naples, Italy

Giorgio Valentini

Università degli Studi di Milano, Dept. of Computer Science, Italy

Giuseppe Palmisano

Department of Biochemistry and Molecular Biology
University of Southern Denmark, Odense M, Denmark

Helen Gika

Chemical Engineering Department, Aristotle University of Thessaloniki, Greece

Hugo Miguel Baptista Carreira dos Santos

REQUIMTE-FCT Universidade NOVA de Lisboa, Portugal

Ignacio Casal

Functional Proteomics Laboratory, Centro de Investigaciones Biológicas (CSIC), Madrid, Spain

Ignacio Ortea

European Commission, Joint Research Center, Institute for Reference Materials and Measurements, Geel, Belgium

Iñaki Álvarez

Institut de Biotecnologia i Biomedicina Vicent Villar Palasí, Universitat Autònoma de Barcelona, Barcelona

Isabel Marcelino

Instituto de Tecnologia Química e Biológica, Oeiras, Portugal

Isabel Liste

Area de Biología Celular y del Desarrollo, Instituto de Salud Carlos III, Madrid, Spain

Isabelle Fournier

University Lille Nord de France, Fundamental & Applied Biological Mass Spectrometry - EA 4550, Villeneuve d'Ascq, France

Jacek Z. Kubiak

CNRS UMR 6061, University of Rennes 1, Institute of Genetics and Development of Rennes, Rennes, France

Jane Thomas-Oates

Centre of Excellence in Mass Spectrometry and Department of Chemistry, University of York, Heslington, UK

Jatin Burniston

Muscle Physiology and Proteomics Laboratory, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Tom Reilly Building, Liverpool, United Kingdom

Jean-Paul Issartel

INSERM U836, Grenoble Institut des Neurosciences, La Tronche, France

Jens Allmer

Molecular Biology and Genetics, Izmir Institute of Technology, Urla, Izmir, Turkey

Jerry Thomas

Technology Facility, Department of Biology, University of York, UK

Jesús Jorrín Novo

Agricultural and Plant Biochemistry, Proteomics Research Group, Department of Biochemistry and Molecular Biology, Córdoba, Spain

Jesus Mateos Martín

Osteoarticular and Aging Research Lab, Proteomics Unit INIBIC-Complejo Hospitalario Universitario de A Coruña, A Coruña, Spain

Joan Cerdà

Laboratory IRTA, Institute of Marine Sciences (CSIC), Passeig marítim 37-49, 08003 Barcelona, Spain

Joan Claria

Department of Biochemistry and Molecular Genetics, Hospital Clínic of Barcelona, Spain

João Rodrigues

Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Portugal

Joaquim ROS

Dept. Ciències Mèdiques Bàsiques. IRB Lleida. University of Lleida, Spain

Joerg Reinders

AG Proteomics, Institute of Functional Genomics, University Regensburg, Germany

Johan Palmfeldt

Research Unit for Molecular Medicine, Aarhus University Hospital, Skejby, Aarhus, Denmark

Jose Andrés Fernández González

Universidad del País Vasco, Facultad de Ciencia y Tecnología, Spain

Jose Câmara

University of Madeira, Funchal, Portugal

Jose Cremata Alvarez

Department of Carbohydrate Chemistry, Center for Genetic Engineering and Biotechnology, Havana, Cuba

Jose Luis Martín-Ventura

IIS-FJD-UAM, Madrid, Spain

José Manuel Bautista

Departamento de Bioquímica y Biología Molecular IV, Universidad Complutense de Madrid, Spain

Jose Manuel Palma

Departamento de Bioquímica, Biología Celular y Molecular de Plantas Estación Experimental del Zaidín, CSIC, Granada, Spain

José Moreira

Danish Center for Translational Breast Cancer Research, Denmark

Juraj Gregan

Max F. Perutz Laboratories, University of Vienna, Austria

Karin Stensjö

Department of Photochemistry and Molecular Science, Ångström laboratory, Uppsala University, Sweden

Kathleen Marchal

CMPG/Bioinformatics, Dep Microbial and Molecular Systems, Leuven, Germany

Kay Ohlendieck

Department of Biology, National University of Ireland, Maynooth, Co. Kildare, Ireland

Keiryn Bennett

CeMM - Center for Molecular Medicine of the Austrian Academy of Sciences Vienna, Austria

Kjell Sergeant

Centre de Recherche Public-Gabriel Lippmann, Department 'Environment and Agro-biotechnologies' (EVA), Luxembourg

Konstantinos Kouremenos

Department of Chemistry, Umea University, Sweden

Lennart Martens

Department of Medical Protein Research, VIB and Department of Biochemistry, Ghent University, Belgium

- Luis P. Fonseca**
Instituto Superior Técnico, Centro de Engenharia Biológica e Química,
Institute for Biotechnology and Bioengineering, Lisboa, Portugal
- Luisa Brito**
Laboratório de Microbiologia, Instituto Superior de Agronomia, Tapada da
Ajuda, Lisbon, Portugal
- Luisa Mannina**
CNR, Istituto di Metodologie Chimiche, Rome, Italy
- Manuel Avilés Sanchez**
Department of Cell Biology and Histology, School of Medicine, University of
Murcia, Spain
- Mar Vilanova**
Misión Biológica de Galicia, Consejo Superior de Investigaciones Científicas,
Pontevedra, Spain
- Marcello Donini**
ENEA -Casaccia Research Center, UTBIORAD-FARM, Biotechnology
Laboratory, Italy
- Marco Lemos**
GIRM & ESTM - Polytechnic Institute of Leiria, Peniche, Portugal
- Marcus Mau**
King's College London, UK
- María Álava**
Departamento de Bioquímica y Biología Molecular y Celular, Facultad de
Ciencias, Universidad de Zaragoza, Spain
Maria De Angelis
Department of Soil, Plant and Food Science, University of Bari Aldo Moro,
Italy
- María de la Fuente**
Legume group, Genetic Resources, Mision Biologica de Galicia-CSIC,
Pontevedra, Spain
- Maria M. Malagón**
Department of Cell Biology, Physiology and Immunology, IMIBIC,
Universidad de Córdoba, Spain
- Maria Gabriela Rivas**
REQUIMTE/CQFB, Departamento de Química, Faculdade de Ciências e
Tecnologia, Universidade Nova de Lisboa, Portugal
- María Mayán**
INIBIC, LaCoruña, Spain
- María Páez de la Cadena**
Department of Biochemistry, Genetics and Immunology, University of Vigo,
Spain
- Marie Arul**
Muséum National Histoire Naturelle, Département RDDM, Plateforme de
spectrométrie de masse et de protéomique, Paris, France
- Marie-Pierre Bousquet**
Institut de Pharmacologie et de Biologie Structurale, UPS/CNRS, Toulouse,
France
- Mario Diniz**
Dept. Química-REQUIMTE, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, Portugal
Mark Davey
Catholic University of Leuven (KU Leuven), Belgium
- Marko Radulovic**
Institute for Oncology and Radiology, Laboratory of Cancer Cell biology,
Belgrade, Serbia
- Martin Hajdich**
Department of Reproduction and Developmental Biology, Institute of Plant
Genetics and Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia
- Martin Kussmann**
Faculty of Science, Aarhus University, Aarhus, Denmark
- Martina Marchetti-Deschmann**
Institute of Chemical Technologies and Analytics, Vienna University of
Technology, Vienna, Austria
- Martine Morzel**
INRA, Centre des Sciences du Goût et de l'Alimentation (CSGA), Dijon,
France
- Maxence Wisztorski**
University Lille 1, Laboratoire de Spectrométrie de Masse Biologique,
Fondamentale & Appliquée, Villeneuve d'Ascq, France
- Meri Hovsepyan**
Institute of Molecular Biology of Armenian National Academy of Sciences
Yerevan, Armenia
- Michalis Nikolaidis**
Department of Physical Education and Sports Science at Serres, Aristotle
University of Thessaloniki, Greece
- Michel Jaquinod**
Exploring the Dynamics of Proteomes/Laboratoire Biologie à Grande Echelle,
Institut de Recherches en Technologies et Sciences pour le Vivant, Grenoble,
France
- Michel Salzet**
Laboratoire de Spectrométrie de Masse Biologique Fondamentale et
Appliquée, INSERM, Villeneuve d'Ascq, France
- Miguel Reboiro Jato**
Escuela Superior de Ingeniería Informática, Ourense, Spain
- Moncef Mrabet**
Laboratory of Legumes (LL), Centre of Biotechnology of Borj-Cédria (CBCB),
Hammam-Lif, Tunisia
- Mónica Botelho**
Centre for the study of animal sciences (CECA)/ICETA, Porto, Portugal
- Monica Carrera**
Institute of Molecular Systems Biology, Zurich, Germany
- Okay Saydam**
Molecular Oncology Laboratory, Division of Neuro-Oncology, Department of
Pediatrics Medical University of Vienna, Austria
- Ola Söderberg**
Department of Immunology, Genetics and Pathology, Uppsala University,
Sweden
- Paloma Sánchez-Bel**
Dpto. Biología del estrés y Patología vegetal, CEBAS-CSIC, Murcia, Spain
- Pantelis Bagos**
Department of Computer Science and Biomedical Informatics, University of
Central Greece, Greece
- Paolo Destefanis**
Department of Urology, "San Giovanni Battista - Molinette" Hospital, Turin,
Italy
- Pasquale Vito**
Università del Sannio, Benevento, Italy
- Patrice Francois**
Genomic Research Laboratory, Service of Infectious Diseases, Department of
Internal Medicine, Geneva
- Patrícia Alexandra Curado Quintas Dinis Poeta**
University of Trás-os-Montes and Alto Douro (UTAD), School of Agrary and
Veterinary Sciences, Veterinary, Science Department, Portugal
- Paul Cutler**
F Hoffman La Roche, Basel, Switzerland
- Paulo Vale**
IPMA - Instituto Português do Mar e da Atmosfera, Lisboa, Portugal
- Pedro Baptista**
Centre for Research in Human Molecular Genetics, Department of
LifeSciences, Faculdade de Ciências e Tecnologia, Universidade Nova de
Lisboa, Caparica, Portugal
- Pedro Rodrigues**
Centro de Ciências do Mar do Algarve, CCMAR, Faro, Portugal
- Pedro Santos**

CBMA-Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, Braga, Portugal

Pedro S. Lazo
Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología del Principado de Asturias (IUOPA), Universidad de Oviedo, Spain

Per Bruheim
Department of Biotechnology, Norwegian University of Science and Technology, Trondheim, Norway

Phillip Cash
Division of Applied Medicine, University of Aberdeen, Scotland

Philipp Hess
Institut Universitaire Mer et Littoral (CNRS - Université de Nantes - Ifremer), Nantes, France

Philippe Castagnone-Sereno
Interactions Biotiques et Sante Vegetale, Sophia Antipolis cedex, France

Pierscionek Barbara
School of Biomedical Sciences, University of Ulster, Cromore Road, Coleraine, BT52 1SA, United Kingdom

Pieter de Lange
Dipartimento di Scienze della Vita, Seconda Università degli Studi di Napoli, Caserta, Italy

Qi Zhu
Dept. Electrical Engineering, ESAT/SCD, Katholieke Universiteit Leuven, Heverlee, Belgium

Ralph Fingerhut
University Children's Hospital, Swiss Newborn Screening Laboratory, Children's Research Center, Zürich, Switzerland

Ralf Hoffmann
Institute of Bioanalytical Chemistry, Center for Biotechnology and Biomedicine, Faculty of Chemistry and Mineralogy, Leipzig University, Germany

Rawi Ramautar
Leiden/Amsterdam Center for Drug Research, Leiden University, The Netherlands

Ricardo Gutiérrez Gallego
Bioanalysis Group, Neuropsychopharmacology Program IMIM-Hospital del Mar & Department of Experimental and Health Sciences, University Pompeu Fabra, Spain

Roman Zubarev
Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

Roque Bru Martinez
Plant Proteomics and Functional Genomics Group, Department of Agrochemistry and Biochemistry, Faculty of Sciences, Alicante University, Spain

Rubén Armañanzas
Computational Intelligence Group, Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Spain

Ruddy Wattiez
Department of Proteomics and Microbiology, University of Mons (UMONS), Belgium

Rune Matthiesen
Institute of Molecular Pathology and Immunology, University of Porto, Portugal

Ruth Birner-Gruenberger
Medical University Graz, Austria

Sabine Luthje
University of Hamburg, Biocenter Klein Flottbek, Hamburg, Germany

Sadin Özdemir
Department of Biology, Faculty of Science and Arts, Siirt University, Turkey

Salvador Ventura
Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Spain

Sandra Kraljevic-Pavelic
University of Rijeka, Department of Biotechnology, Croatia

Sebastian Galuska
Institute of Biochemistry, Faculty of Medicine, Justus-Liebig-University of Giessen, Germany

Serge Cosnier
Department of Molecular Chemistry, Grenoble university/CNRS, Grenoble, France

Serhat Döker
Cankiri Karatekin University, Chemistry Department, Cankiri, Turkey

Shan He
Centre for Systems Biology, School of Biosciences and School of Computer Science, University of Birmingham, England

Silvia Mazzuca
Plan Cell Physiology Laboratory, Department of Ecology, University of Calabria, Italy

Simona Martinotti
Dipartimento di Scienze e Innovazione Tecnologica, DiSIT, University of Piemonte Orientale, Alessandria, Italy

Soile Tapio
Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Radiation Biology, Neuherberg, Germany

Sophia Kossida
Biomedical Research Foundation, Academy of Athens, Department of Biotechnology, Athens, Greece

Spiros D. Garbis
Biomedical Research Foundation of the Academy of Athens, Center for Basic Research - Division of Biotechnology, Greece

Steve Thany
Laboratoire Récepteurs et Canaux Ioniques Membranaires, UFR Science, Université d'Angers, France

Stefania Orrù
University of Naples Parthenope, Naples, Italy

Stefanie Hauck
Research Unit Protein Science, Helmholtz Center Munich, Neuherberg, Germany

Stefano Curcio
Department of Engineering Modeling, Laboratory of Transport Phenomena and Biotechnology University of Calabria, Italy

Susana Cristóbal
Department of Clinical and Experimental Medicine Faculty of Health Science Linköping University, Sweden

Tàmara García Barrera
Departamento de Química y Ciencia de los Materiales, Facultad de Ciencias Experimentales, Universidad de Huelva, Spain

Theodore Alexandrov
University of Bremen, Center for Industrial Mathematics, Germany

Thole Züchner
Ultrasensitive Protein Detection Unit, Leipzig University, Center for Biotechnology and Biomedicine, Institute of Bioanalytical Chemistry, Germany

Tiziana Bonaldi
Department of Experimental Oncology, European Institute of Oncology, Via Adamello 16, 20139 Milan, Italy

Tomris Ozben
Akdeniz University Medical Faculty Department of Clinical Biochemistry, Antalya, Turkey

Tsangaris George
Proteomics Research Unit, Center of Basic Research II Foundation of Biomedical Research of the Academy of Athens, Greece

Üner Kolukisaoglu

Center for Plant Molecular Biology, Eberhard-Karls University Tübingen, Tübingen, Germany

Valeria Bertagnolo

Department of Morphology and Embryology University of Ferrara, Italy

Vera Muccilli

Dipartimento di Scienze Chimiche, Università di Catania, Catania, Italy

Veronica Mainini

Dept. Health Science, University of Milano-Bicocca, Faculty of Medicine, Monza (MB), Italy

Vicenta Martínez-Zorzano

Department of Biochemistry, Genetics and Immunology University of Vigo, Spain

Virginie Brun

French Atomic Energy Commission and *French National Institute for Health*

and Medical Research, France

Vittoria Matafora

Biological Mass Spectrometry Unit, San Raffaele Scientific Institute, Milan, Italy

Vladislav Khrustalev

Department of General Chemistry, Belarussian, State Medical University, Dzerzhinskogo, Minsk, Belarus

Xiaoze Zhang

Department of Medicine, University of Fribourg, Switzerland

Yuri van der Burgt

Leiden University Medical Center, Department of Parasitology, The Netherlands

SOUTH AMERICA

Alessandro Farias

Neuroimmunomodulation Group, department of Genetics, Evolution and Bioagents, University of Campinas - SP - Brazil

Alexandra Sawaya

Department of Plant Biology, Institute of Biology, UNICAMP, Campinas, São Paulo, Brazil

Andréa P.B. Gollucke

Hexalab/Catholic University of Santos, Brazil

Arlindo Moura

Department of Animal Science - College of Agricultural Sciences - Federal University of Ceara, Fortaleza, Brasil

Bruno Lomonte

Instituto Clodomiro Picado, Universidad de Costa Rica

Deborah Schechtman

Department of Biochemistry, Chemistry Institute, University of São Paulo, Brazil

Edson Guimarães Lo Turco

São Paulo Federal University, Brasil

Elisabeth Schwartz

Department of Physiological Sciences, Institute of Biological Sciences, University of Brasilia, Brazil

Fabio Ribeiro Cerqueira

Department of Informatics and NuBio (Research Group for Bioinformatics), University of Vicoso, Brazil

Fernando Barbosa

Faculty of Pharmaceutical Sciences of Ribeirão Preto University of São Paulo, Brazil

Hugo Eduardo Cerecetto

Grupo de Química Medicinal, Facultad de Química, Universidad de la República, Montevideo, Uruguay

Luis Pacheco

Institute of Health Sciences, Federal University of Bahia, Salvador, Brazil

Mário Hiroyuki Hirata

Laboratório de Biologia Molecular Aplicado ao Diagnóstico, Departamento de Análises Clínicas e Toxicológicas, Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Brazil

Jan Schripsema

Grupo Metabolômica, Laboratório de Ciências Químicas, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Brazil

Jorg Kobarg

Centro Nacional de Pesquisa em Energia e Materiais, Laboratório Nacional de Biociências, Brazil

Marcelo Bento Soares

Cancer Biology and Epigenomics Program, Children's Memorial Research Center, Professor of Pediatrics, Northwestern University's Feinberg School of Medicine

Mario Palma

Center of Study of Social Insects (CEIS)/Dept. Biology, Institute of Biosciences, University of São Paulo State (UNESP), Rio Claro - SP Brazil

Rinaldo Wellerson Pereira

Programa de Pós Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brazil

Roberto Bobadilla

BioSigma S.A., Santiago de Chile, Chile

Rossana Arroyo

Department of Infectomic and Molecular Biology, Center of Research and Advanced Studies of the National, Polytechnical Institute (CINVESTAV-IPN), Mexico City, Mexico

Rubem Menna Barreto

Laboratorio de Biología Celular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Vasco Azevedo

Biological Sciences Institute, Federal University of Minas Gerais, Brazil

NORTH AMERICA

Adam Vigil

University of California, Irvine, USA

Akeel Baig

Hoffmann-La Roche Limited, Pharma Research Toronto, Toronto, Ontario, Canada

Alexander Statnikov

Center for Health Informatics and Bioinformatics, New York University School of Medicine, New York

Amosy M'Koma

Vanderbilt University School of Medicine, Department of General Surgery,

Colon and Rectal Surgery, Nashville, USA

Amrita Cheema

Georgetown Lombardi Comprehensive Cancer Center, USA

Anthony Gramolini

Department of Physiology, Faculty of Medicine, University of Toronto, Canada

Anas Abdel Rahman

Department of Chemistry, Memorial University of Newfoundland and Labrador St. John's, Canada

Christina Ferreira

Purdue University - Aston Laboratories of Mass Spectrometry, Hall for Discovery and Learning Research, West Lafayette, US

Christoph Borchert
Biochemistry & Microbiology, University of Victoria, UVic Genome British Columbia Proteomics Centre, Canada

Dajana Vuckovic
University of Toronto, Donnelly Centre for Cellular + Biomolecular Research, Canada

David Gibson
University of Colorado Denver, Anschutz Medical Campus, Division of Endocrinology, Metabolism and Diabetes, Aurora, USA

Deyu Xie
Department of Plant Biology, Raleigh, USA

Edgar Jaimes
University of Alabama at Birmingham, USA

Eric McLamore
University of Florida, Agricultural & Biological Engineering, Gainesville, USA

Eustache Paramithiotis
Caprion Proteomics Inc., Montreal, Canada

FangXiang Wu
University of Saskatchewan, Saskatoon, Canada

Fouad Daayf
Department of Plant Science, University of Manitoba, Winnipeg, Manitoba, Canada

Haitao Lu
Washington University School of Medicine, Saint Louis, USA

Hexin Chen
University of South Carolina, Columbia, USA

Hsiao-Ching Liu
232D Polk Hall, Department of Animal Science, North Carolina State University Raleigh, USA

Hui Zhang
Johns Hopkins University, MD, USA

Ing-Feng Chang
Institute of Plant Biology, National Taiwan University, Taipei, Taiwan

Irwin Kurland
Albert Einstein College of Medicine, Associate Professor, Dept of Medicine, USA

Jagjit Yadav
Microbial Pathogenesis and Toxicogenomics, Laboratory, Environmental Genetics and Molecular, Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Ohio, USA

Jianbo Yao
Division of Animal and Nutritional Sciences, USA

Jiaxu Li
Department of Biochemistry and Molecular Biology, Mississippi State University, USA

Jiping Zhu
Exposure and Biomonitoring Division, Health Canada, Ottawa, Canada

Jiri Adamec
Department of Biochemistry & Redox Biology Center, University of Nebraska, Lincoln Nebraska, USA

Jiye Ai
University of California, Los Angeles

John McLean
Department of Chemistry, Vanderbilt University, Nashville, TN, USA

Joshua Heazlewood
Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Kenneth Yu
Memorial Sloan Kettering Cancer Center, New York, USA

Laszlo Prokai
Department of Molecular Biology & Immunology, University of North Texas Health Science Center, Fort Worth, USA

Lei Li
University of Virginia, USA

Leonard Foster
Centre for High-throughput Biology, University of British Columbia, Vancouver, BC, Canada

Madhulika Gupta
Children's Health Research Institute, University of Western Ontario London, ON, Canada

Masaru Miyagi
Case Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, USA

Michael H.A. Roehrl
Department of Pathology and Laboratory Medicine, Boston Medical Center Boston, USA

Ming Zhan
National Institute on Aging, Maryland, USA

Nicholas Seyfried
Emory University School of Medicine, Atlanta, USA

Olgica Trenchevska
Molecular Biomarkers, Biodesign Institute at Arizona State University, USA

Peter Nemes
US Food and Drug Administration (FDA), Silver Spring, USA

R. John Solaro
University of Illinois College of Medicine, USA

Rabih Jabbour
Science Application International Corporation, Maryland, USA

Ramesh Katam
Plant Biotechnology Lab, Florida A and M University, FL, USA

Robert L. Hettich
Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, USA

Robert Powers
University of Nebraska-Lincoln, Department of Chemistry, USA

Shen S. Hu
UCLA School of Dentistry, Dental Research Institute, UCLA Jonsson Comprehensive Cancer Center, Los Angeles CA, USA

Shiva M. Singh
University of Western Ontario, Canada

Susan Hester
United States Environmental Protection Agency, Durnam, USA

Terry D. Cyr
Genomics Laboratories, Centre for Vaccine Evaluation, Biologics and Genetic Therapies Directorate, Health Products and Foods Branch, Health Canada, Ontario, Canada

Thibault Mayor
Department of Biochemistry and Molecular Biology, Centre for High-Throughput Biology (CHiBi), University of British Columbia, Canada

Thomas Conrads
USA

Thomas Kislinger
Department of Medical Biophysics, University of Toronto, Canada

Wan Jin Jahng
Department of Biological Sciences, Michigan Technological University, USA

Wayne Zhou
Marine Biology Laboratory, Woods Hole, MA, USA

Wei Jia
US Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Wei-Jun Qian
Pacific Northwest National Laboratory, USA

William A LaFramboise
Department of Pathology, University of Pittsburgh School of Medicine

Shadyside Hospital, Pittsburgh, USA

Xiangjia Min

Center for Applied Chemical Biology, Department of Biological Sciences
Youngstown State University, USA

Xiaoyan Jiang

Senior Scientist, Terry Fox Laboratory, BC Cancer Agency, Vancouver,
Canada

Xu-Liang Cao

Food Research Division, Bureau of Chemical Safety, Health Canada, Ottawa,
Canada

Xuequn Chen

Department of Molecular & Integrative Physiology, University of Michigan,
Ann Arbor, USA

Ye Fang

Biochemical Technologies, Science and Technology Division, Corning
Incorporated, USA

Ying Qu

Microdialysis Experts Consultant Service, San Diego, USA

Ying Xu

Department of Biochemistry and Molecular Biology, Institute of
Bioinformatics, University of Georgia, Life Sciences Building
Athens, GA, USA

JOURNAL OF INTEGRATED OMICS

A methodological Journal

CONTENTS OF VOLUME 2 | ISSUE 2 | DECEMBER 2012

LETTER TO THE EDITOR

- A Call for Benchmark Data in Mass Spectrometry-Based Proteomics. 1
Jens Allmer.

REVIEW ARTICLES

- Integrated Omics Analysis of Sjogren's Syndrome. 6
Jiye Ai, Sizhe Feng, Kaori Misuno, Shen Hu.
- A survey on coronary heart disease related signal pathways, drug targets and pharmacological interventions. 11
Peng Jiang, Runhui Liu, Weidong Zhang.
- Organelle proteomics in skeletal muscle biology. 27
Kay Ohlendieck.

ORIGINAL ARTICLES

- Protein thiols as novel biomarkers in ecotoxicology: A case study of oxidative stress in *Mytilus edulis* sampled near a former industrial site in Cork Harbour, Ireland. 39
Sara Tedesco, Siti Nur Tahirah Jaafar, Ana Varela Coelho and David Sheehan.
- Identification of outer membrane proteins of *Edwardsiella tarda* in response to high concentration of copper. 48
Chao Wang, Haili Zhang, Hui Li, Xuan-xian Peng.
- Comparative immunoproteome analysis of the response of susceptible A.BY/SnJ and resistant C57BL/6 mice to Coxsackievirus B3-infection. 54
Elke Hammer, Truong Quoc Phong, Leif Steil, Manuela Gesell Salazar, Christian Scharf, Reinhard Kandolf, Stephan B. Felix, Heyo K. Kroemer, Karin Klingel, Uwe Völker.
- Secretome differences between the taxonomically related but clinically differing mycobacterial species *Mycobacterium abscessus* and *M. chelonae*. 64
Jagjit S. Yadav and Manish Gupta.
- A robust permutation test for quantitative SILAC proteomics experiments. 80
Hien D. Nguyen, Ian A. Wood, Michelle M. Hill.
- Analysis of the rat primary hepatocyte nuclear proteome through sub-cellular fractionation. 94

Cliff Rowe, Roz E. Jenkins, Neil R. Kitteringham, B. Kevin Park, Christopher E.P. Goldring.

New insights in *Trypanosoma cruzi* proteomic map: further post-translational modifications and potential drug targets in Y strain epimastigotes. 106

Daniela Gois Beghini, André Teixeira da Silva Ferreira, Vivian Corrêa de Almeida, Marcelle Almeida Caminha, Floriano Paes Silva-Jr, Jonas Perales, Rubem Figueiredo Sadok Menna-Barreto.

Normalization of protein at different stages in SILAC subcellular proteomics affects functional analysis. 114

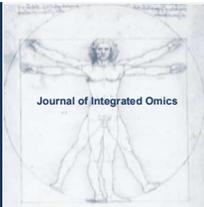
Kerry L. Inder, Dorothy Loo, Yu Zi Zheng, Robert G. Parton, Leonard J. Foster, Michelle M. Hill.

Proteomic analysis of bronchoalveolar lavage fluid in an equine model of asthma during a natural antigen exposure trial. 123

Marybeth Miskovic Feutz, C. Paige Riley, Xiang Zhang, Jiri Adamec, Craig Thompson, Laurent L. Couetil.

Electrophysiological and Proteomic Studies of *Protobothrops mangshensis* Venom Revealed Its High Bioactivities and Toxicities. 132

Yu Liu, Xiaojuan Wang, Zhe Wu, Ping Chen, Xiang Yang, Dongsong Nie, Jianmin Yi, Jizu Yi and Zhonghua Liu.



LETTER TO THE EDITOR | DOI: 10.5584/jiomics.v2i2.113

A Call for Benchmark Data in Mass Spectrometry-Based Proteomics

Jens Allmer

Molecular Biology and Genetics, Izmir Institute of Technology, Urla, Izmir.

Received: 02 October 2012 Accepted: 26 October 2012 Available Online: 28 October 2012

ABSTRACT

Proteomics is a quickly developing field. New and better mass spectrometers, the platform of choice in proteomics, are being introduced frequently. New algorithms for the analysis of mass spectrometric data and assignment of amino acid sequence to tandem mass spectra are also presented on a frequent basis. Unfortunately, the best application area for these algorithms cannot be established at the moment. Furthermore, even the accuracy of the algorithms and their relative performance cannot be established. This is due to the lack of proper benchmark data. This letter first introduces the field of mass spectrometry-based proteomics and then defines the expectations of a well-designed benchmark dataset. Thereafter, the current situation is compared to this ideal. A call for the creation of a proper benchmark dataset is then placed and it is explained how measurement should be performed. Finally, the benefits for the research community are highlighted.

Keywords: Mass spectrometry; Proteomics; Benchmark data; Database search; De novo sequencing; Fragmentation analysis.

Abbreviations:

MS, Mass Spectrometry; MALDI, Matrix Assisted Laser Desorption Ionization; Da, Dalton; m/z, Mass to charge ratio; TOF, Time of Flight; NCBI, National Center for Biotechnology Information.

1. Introduction

Proteomics is the study of the proteome, the entirety of proteins, their spatial and temporal expression patterns, their modifications, interactions, and of course their functions. To elucidate the proteome of any higher eukaryote is currently a futile endeavor but subsets thereof can be investigated. Mass spectrometry (MS) has become the tool of choice in proteomics [1] and is used to establish protein sequence, quantity, and modification.

The output of any mass spectrometer is a list of mass to charge ratios (m/z) of the peptides in a sample. To derive further information additional stages of MS can be employed following a fragmentation of the peptide precursor for which many methods are available [2]. Tandem-MS (MS/MS, MS²) spectra contain a list of m/z values of the fragmented pep-

tide.

There are basically two ways to assign a peptide sequence to an MS/MS spectrum from an unknown precursor. One method, termed database search, depends on the availability of a database of either sequences or reference tandem-MS spectra [3]. In contrast to that, *de novo* sequencing derives the sequence solely from the MS/MS spectrum [2]. For both database search and *de novo* sequencing many algorithms have been developed for computational analysis of MS² spectra.

An abundance of different mass spectrometers are now available, which can further be coupled with a large number of fragmentation methods. This leads to a large number of possible measurement methodologies. According to the 'no

*Corresponding author: Jens Allmer, Phone Number.: 00902327507517, Fax Number: 00902327507509, E-mail address: jens@allmer.de, Postal Address: Assoc. Prof. Dr. Jens Allmer, Molecular Biology and Genetics, Izmir Institute of Technology, Gulbahce Campus, Urla, Izmir, Turkey.

free lunch' theorem, there is no one algorithm which is best for all instances of a problem [4,5] let alone for all problems posed by varying combinations of mass spectrometers and fragmentation methods. This seems to be a problem widely ignored in biological sciences where a tool if it works on a small number of instances of the general problem will be quickly adopted to solve even problems well beyond its domain [6]. In mass spectrometry this can be exemplified by the ubiquitous usage of Mascot [7] for many types of MS measurements while it was initially intended to be used for peptide mass fingerprinting on matrix assisted laser desorption ionization - time of flight (MALDI-TOF) mass spectrometers. Admittedly, Mascot has since been extended to include fragmentation spectra but it is still targeted towards MALDI. It seems clear that any of the algorithms in database search or *de novo* sequencing perform differently on diverse data.

Unfortunately, it is unclear which algorithm is best for which combination of mass spectrometer and fragmentation method. Some studies have investigated the performance of database search algorithms [3,8–10] and other studies the accuracy of *de novo* sequencing algorithms [11–13]. A mere review of the relevant publications is in this case not possible since all new algorithms are usually developed on different mass spectrometric data sets. Therefore, it is essential to employ the algorithms to be compared on the same dataset before comparing their performance.

Only few benchmark datasets have been published in mass spectrometry-based proteomics which would allow such a comparison. Whether these can truly be called benchmark datasets will be investigated below (Section 3), but before that a general description of a proper benchmark dataset will be given (Section 2).

In order to further the field of mass spectrometry and to develop useful algorithms, this letter is also a call to action. Anyone having access to synthetic peptides should participate in the development of a first benchmark dataset and provide a few hundred spectra per synthetic peptide so that a comprehensive benchmark dataset, based in ground truth, can be created.

2. Benchmark Data

In the widest sense of the word a benchmark is a standardized performance test. In mass spectrometry-based proteomics a benchmark would thus measure the performance of algorithms to assign a sequence to an MS/MS spectrum. A benchmark dataset for mass spectrometry-based proteomics must thus consist of MS/MS spectra and their correct sequence annotation. Additionally, the mass spectrometer used, the fragmentation method, and the measurement settings should be specified.

Aniba and colleagues defined six measures for well-constructed benchmark datasets [14] which will be discussed in respect to mass spectrometry-based proteomics in the following.

2.1. Relevance

In mass spectrometry-based proteomics and in accordance with the 'no free lunch' theorem, a benchmark dataset could target a combination of a particular mass spectrometer and fragmentation method. This platform should be used to generate enough spectra from enough different peptides so that the scope of the benchmark dataset can be fulfilled. In practice that means that if the benchmark dataset is targeted to test the performance of database search algorithms, it should consist of spectra similar to the ones that may be expected in experimental studies. For mass spectrometry this is difficult since the peptide sequence seems to have a strong influence on fragmentation [15] and therefore the resulting spectrum. Therefore, measurements from all MS platforms are needed.

2.2. Solvability

The benchmark dataset needs to be solvable. It should be neither too hard nor too easy for existing algorithms so that the benchmark can be used to differentiate performance among algorithms. In practice this disqualifies all existing benchmark datasets since all of them present the same "chicken or egg" problem: The MS/MS spectra are usually derived from a protein digest and are then assigned sequences using a database search algorithm. Unfortunately, in many cases algorithms do not agree on the sequence and thus the correctness of the sequence cannot be guaranteed. Since in existing datasets this is unaccounted for, they are not correctly solvable as they are not correct in themselves. Hence, current algorithms are presented with an insurmountable challenge.

2.3. Scalability

A benchmark dataset should present problems at various level of difficulty so that it can scale with the maturity of algorithms. Current datasets might contain examples of different difficulty, but they are not properly annotated and thus not applicable for benchmarking. In mass spectrometry, datasets are exceedingly difficult to solve and more simple datasets like repeated measurements of synthetic peptides are needed to present less perplexing problems and establish the accuracy of the foundation of more advanced methodologies. Later, problems like larger tolerances in the m/z measurement, increasing noise level (i.e.: unexplained peaks from currently poorly understood fragmentation pathways such as sequence scrambling [16], and precursor ions of higher charge can be addressed. Peptides, which lead to fewer fragments than expected, can present a difficult challenge for algorithms following successes on simple benchmark datasets. Even more difficulty can be created by measuring co-eluting and co-fragmenting peptides, somewhat rare, but yet significant problems.

2.4. Accessibility

The benchmark dataset needs to be accessible so that algorithm developers or users can benchmark the tools they develop or want to use. Mass spectrometric data may have an intrinsic problem since proper benchmark datasets will likely have large file sizes (several gigabytes) which may be difficult to host and/or transfer. Nonetheless, several platforms for data sharing have been established (see Section 3).

2.5. Independence

Aniba and colleagues [14] make a very important point: “The methods or approaches to be evaluated should not be used to construct the gold standard tests. Otherwise, the developers could be accused of ‘cheating’, i.e. designing the benchmark to suit the software”. This seems to be obvious but many new algorithms in mass spectrometry-based proteomics are published alongside with their own datasets. This is obviously due to the fact that no proper benchmark dataset is available and should not be used to accuse developers of cheating in this case.

2.6. Evolution

The benchmark dataset needs to be modified constantly to prevent researchers from optimizing their algorithms to solve it. New suitable datasets should thus be published frequently to test the performance of existing algorithms. The need for scalability must be taken into account and subsequent datasets should be gradually more challenging.

Currently, none of the available mass spectrometric datasets adhere to more than one of these six requirements. Nonetheless, they are being used as benchmark datasets with all the negative consequences that entail such as low or unknown accuracy of commonly used algorithms.

3. Mass Spectral Data

Although it is not mandatory for most journals publishing in the area of proteomics, some journals and funders make it mandatory that raw data be made publicly available [17]. Currently, not enough is being done to ensure that raw data is made available and a suitable incentive for researchers to comply is yet to be found [18]. Nonetheless, an abundance of mass spectrometric measurements have been made available in a number of public repositories. The major repositories, in no particular order, are Global Proteome Machine Database [19], PeptideAtlas [20], Proteomics IDentifications Database [21], Proteome Commons’ Tranche (<https://proteomecommons.org/tranche/>), and NCBI’s Peptidome [22]. Other smaller or more targeted collections are also available and have been reviewed in Mead and colleagues [23] and Riffle and Eng [24].

Proteomics repositories provide large amounts of raw mass spectrometric data which may also be annotated through database search and may be useful for research [17].

However, there is a hidden “chicken or egg” problem present. MS/MS spectra are usually identified using the database search engine of choice of the laboratory that made the measurements. And these assignments are then used to train new algorithms. An example for this is the benchmark dataset created by Keller and colleagues who used Sequest [25] to assign a sequence to the tandem-MS spectra [26]. Later it was shown that the data set contains additional possible assignments [27] which were not given in the initial publication. It is also likely that many assignments for the Keller et al. dataset are wrong since we were not able to reproduce them with other database search engines or *de novo* sequencing tools (data not shown). This leads to the serious problem that new algorithms are trained on a dataset with assignments of another algorithm. Thus all new trained algorithms will likely duplicate errors done by the algorithm used to assign peptides to spectra during the creation of the dataset.

The field of mass spectrometry-based proteomics is large and interfaces with many instruments other than just mass spectrometers. This fact is mirrored in the approaches used to develop benchmark datasets. For example the Keller et al. dataset closely mirrors standard high throughput studies. Two more recent datasets do the same and although they are more elaborate still present the same “chicken or egg” problem. Wessels et al. present a very comprehensive dataset based on the *Escherichia coli* proteome with additional spiked in known protein digests as a real life challenge [28]. Beasley-Green and colleagues also chose a model organism (*Saccharomyces cerevisiae*) to design their dataset [29]. Another approach for designing benchmark datasets is based on simulation [30], but this approach, while offering a ground truth, is synthetic and should itself be benchmarked on real data. All mentioned datasets aim to benchmark the overall process and neglect the fact that each individual process should be properly benchmarked before integration testing can be performed. Therefore it is necessary to develop more targeted and well-designed benchmark datasets to prove the effectiveness of all modules of the overall mass spectrometry-based proteomics workflow.

4. Call for Action

As detailed above, there is no publicly available dataset which is solvable and thus at least two of the measures for well-constructed benchmark datasets are violated. It is the aim of this letter to engage the mass spectrometric community in creating a compliant benchmark dataset.

As current datasets are not solvable due to the “chicken or egg” problem, it is necessary to assure the assigned sequence by a different means than any of the current database search or *de novo* sequencing algorithms. The simplest way that the sequence assignment can be guaranteed is by directly injecting/spotting pure synthetic peptides. The resulting dataset will be solvable in theory and thus would enable true benchmarking of current algorithms and would enable developers

to benchmark their new algorithms.

This letter urges anyone with access to synthetic peptides and mass spectrometers to measure them in the following way:

1. Measure the synthetic peptide at high concentration (10-50 MS/MS spectra)
2. Decrease the concentration (e.g.: lower the flow rate) and measure 10-50 MS/MS spectra
3. Keep decreasing the concentration and measure 10-50 MS/MS spectra
4. Stop measuring when the signal is disappearing in the noise.

Please submit the measurements for individual synthetic peptides in mzXML [31] or mzML [32,33] file format to the author. A comprehensive dataset will be created and deposited in Proteomics IDentifications Database [21] and NCBI's Peptidome [22]. For this dataset an additional website will be created, crediting anyone submitting data, making available the data, and providing additional information about the dataset. A preliminary version of this website can be reached at <http://msbenchmark.biolnk.com>.

The dataset is relevant as MS/MS spectra are measured as they would be measured in current experimental procedures although the fact that usually mixtures are investigated is ignored. Additional problem due to liquid chromatography are also ignored. It is the aim to have undisturbed MS/MS measurements of known precursors and varying quality that is solvable in the domain of assigning sequence to MS/MS spectra. As different spectral qualities are created it is to some degree scalable. All benchmark datasets created will be accessible but their relevance will decrease with time when more challenging datasets will be prepared. Thus the dataset will evolve and make optimization targeting the dataset only possible for older benchmarks. Finally, as this is a community effort, the independence of the data is guaranteed. The resulting dataset thus adheres to all features of proper benchmark datasets.

5. Community Benefits

Some of the benefits that the mass spectrometric community can gain from proper benchmark datasets are quite obvious. For experimentalists it would be important to know which algorithms are best for their mass spectrometer and fragmentation method combination. There is no silver bullet among algorithms so no one algorithm can perform best on all measurement platforms. Currently, it is not possible to compare algorithms but given a comparison the best suited computational analysis strategy can be pursued. This in turn leads to more accurate data entering public repositories, again benefiting experimentalists even outside of the field of mass spectrometry-based proteomics. It has also been pointed out by Noble and MacCoss that a critical assessment for algorithms in the field is lacking [34] which further underlines the issue raised above.

Clearly, developers of new database search or *de novo* se-

quencing algorithms will be enabled to find out for which platform their algorithm is best suited and how it compares to the performance of other algorithms on a publicly available well annotated and solvable benchmark dataset presented in a standard format.

The benchmark dataset is not limited to the benchmarking of algorithms; it can also be used in novel ways that cannot be foreseen now. However, some additional benefits are for instance the ability to employ data mining on the benchmark dataset to learn general parameters about for example peptide fragmentation [24]. A large number of synthetic peptides measured with a variety of mass spectrometric platforms will enable theoretical chemists to develop new fragmentation models. This in turn will enhance peptide identification when these models are integrated into database search or *de novo* sequencing algorithms.

Finally, the associated web page which credits laboratories, researchers, and their measurements may foster the exchange of measurements or samples within the resulting community (<http://www.biolnk.com/msbenchmark>).

6. Concluding Remarks

This letter was meant to prove that there is a current lack of benchmark data for assigning sequence to MS/MS spectra. This influences the speed of development of the field of mass spectrometry-based proteomics. Furthermore, it forces experimentalists to work with computational analysis tools of unknown accuracy. The benefits of making benchmark data available have been briefly mentioned. It will help increase the accuracy of sequence assignments and in turn the accuracy of conclusions from experimental data in literature and in public databases.

This letter is a call for the submission of MS/MS measurements of synthetic peptides and intends to create an initiative to develop a first proper benchmark for the field of mass spectrometry-based proteomics.

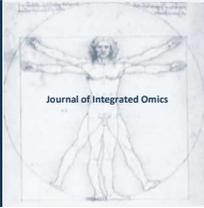
Acknowledgements

This work was in part supported by an award received from the Turkish Academy of Sciences (TÜBA, www.tuba.gov.tr).

References

1. M. Mann, R.C.C. Hendrickson, A. Pandey, Annual Review of Biochemistry 70 (2001) 437–73. DOI: 10.1146/annurev.biochem.70.1.437
2. J. Allmer, Expert Review of Proteomics 8 (2011) 645–57. DOI: 10.1586/epr.11.54
3. E. Kapp, F. Schütz, Current Protocols in Protein Science / Editorial Board, John E. Coligan ... [et Al.] Chapter 25 (2007) Unit25.2. DOI: 10.1002/0471140864.ps2502s49
4. D.H. Wolpert, W.G. Macready, IEEE Transactions on Evo-

- lutionary Computation 1 (1995) 67–82.
5. H.P. Schwefel, in: T. Back, D.B. Fogel, Z. Michalevicz (Eds.), *Evolutionary Computation 1 Basic Algorithms and Operators*, Institute of Physics Publishing, Bristol and Philadelphia, 2000, pp. 20–22.
 6. J. Brownlee, *On Biologically Inspired Computation*, 2005.
 7. D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, *Electrophoresis* 20 (1999) 3551–67. DOI: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
 8. E.A. Kapp, F. Schutz, L.M. Connolly, J.A. Chakel, J.E. Meza, C.A. Miller, D. Fenyo, J.K. Eng, J.N. Adkins, G.S. Omenn, R.J. Simpson, *Proteomics* 5 (2005) 3475–90.
 9. D.C. Chamrad, G. Korting, K. Stuhler, H.E. Meyer, J. Klose, M. Bluggel, *Proteomics* 4 (2004) 619–28.
 10. I. Shadforth, D. Crowther, C. Bessant, *Proteomics* 5 (2005) 4082–95. DOI: 10.1002/pmic.200402091
 11. S. Bringans, T.S. Kendrick, J. Lui, R. Lipscombe, *Rapid Communications in Mass Spectrometry* 22 (2008) 3450–3454. DOI: 10.1002/rcm.3752
 12. S. Pevtsov, I. Fedulova, H. Mirzaei, C. Buck, X. Zhang, *J Proteome Res* 5 (2006) 3018–28.
 13. E. Pitzer, A. Masselot, J. Colinge, *Proteomics* 7 (2007) 3051–4. DOI: 10.1002/pmic.200700224
 14. M.R. Aniba, O. Poch, J.D. Thompson, *Nucleic Acids Research* 38 (2010) 7353–63. DOI: 10.1093/nar/gkq625
 15. A.R. Dongre, J.L. Jones, A. Somogyi, V.H. Wysocki, *Journal of the American Chemical Society* 118 (1995) 8365–8374.
 16. A.E. Atik, T. Yalcin, *Journal of the American Society for Mass Spectrometry* 22 (2011) 38–48. DOI: 10.1007/s13361-010-0018-3
 17. J.A. Vizcaíno, J.M. Foster, L. Martens, *Journal of Proteomics* 73 (2010) 2136–46. DOI: 10.1016/j.jprot.2010.06.008
 18. Editorial, *Nature Biotechnology* 27 (2009) 579. DOI: 10.1038/nbt0709-579
 19. R. Craig, J.P. Cortens, R.C. Beavis, *Journal of Proteome Research* 3 (2004) 1234–42. DOI: 10.1021/pr049882h
 20. E.W. Deutsch, *Methods in Molecular Biology* (Clifton, N.J.) 604 (2010) 285–96. DOI: 10.1007/978-1-60761-444-9_19
 21. J.A. Vizcaíno, R. Côté, F. Reisinger, H. Barsnes, J.M. Foster, J. Rameseder, H. Hermjakob, L. Martens, *Nucleic Acids Research* 38 (2010) D736–42. DOI: 10.1093/nar/gkp964
 22. L. Ji, T. Barrett, O. Ayanbule, D.B. Troup, D. Rudnev, R.N. Muertter, M. Tomashevsky, A. Soboleva, D.J. Slotta, *Nucleic Acids Research* 38 (2010) D731–5. DOI: 10.1093/nar/gkp1047
 23. J.A. Mead, L. Bianco, C. Bessant, *Proteomics* 9 (2009) 861–81. DOI: 10.1002/pmic.200800553
 24. M. Riffle, J.K. Eng, *Proteomics* 9 (2009) 4653–63. DOI: 10.1002/pmic.200900216
 25. J. Eng, A.L. McCormack, J.R. Yates, *J Am Soc Mass Spectrom* 5 (1994) 976–989.
 26. A. Keller, S. Purvine, A.I. Nesvizhskii, S. Stolyar, D.R. Goodlett, E. Kolker, *Omics* 6 (2002) 207–12.
 27. D. Tsur, S. Tanner, E. Zandi, V. Bafna, P.A.P. a Pevzner, *Nature Biotechnology* 23 (2005) 1562–7. DOI: 10.1038/nbt1168
 28. H.J.C.T. Wessels, T.G. Bloemberg, M. van Dael, R. Wehrens, L.M.C. Buydens, L.P. van den Heuvel, J. Gloerich, *Proteomics* 12 (2012) 2276–81. DOI: 10.1002/pmic.201100284
 29. A. Beasley-Green, D. Bunk, P. Rudnick, L. Kilpatrick, K. Phinney, *Proteomics* 12 (2012) 923–31. DOI: 10.1002/pmic.201100522
 30. C. Bielow, S. Aiche, S. Andreotti, K. Reinert, *Journal of Proteome Research* 10 (2011) 2922–9. DOI: 10.1021/pr200155f
 31. P.G.A. Pedrioli, J.K. Eng, R. Hubley, M. Vogelzang, E.W. Deutsch, B. Raught, B. Pratt, E. Nilsson, R.H. Angeletti, R. Apweiler, K. Cheung, C.E. Costello, H. Hermjakob, S. Huang, R.K. Julian, E. Kapp, M.E. McComb, S.G. Oliver, G. Omenn, N.W. Paton, R. Simpson, R. Smith, C.F. Taylor, W. Zhu, R. Aebersold, *Nature Biotechnology* 22 (2004) 1459–66.
 32. E.W. Deutsch, *Methods in Molecular Biology* (Clifton, N.J.) 604 (2010) 319–331. DOI: 10.1007/978-1-60761-444-9_22
 33. L. Martens, M. Chambers, M. Sturm, D. Kessner, F. Levander, J. Shofstahl, W.H. Tang, A. Römpf, S. Neumann, A.D. Pizarro, L. Montecchi-Palazzi, N. Tasman, M. Coleman, F. Reisinger, P. Souda, H. Hermjakob, P.-A. Binz, E.W. Deutsch, *Molecular & Cellular Proteomics: MCP* 10 (2011) R110.000133. DOI: 10.1074/mcp.R110.000133
 34. W.S. Noble, M.J. MacCoss, *PLoS Computational Biology* 8 (2012) e1002296. DOI: 10.1371/journal.pcbi.1002296



JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM



REVIEW ARTICLE | DOI: 10.5584/jiomics.v2i2.97

Integrated Omics Analysis of Sjögren's Syndrome

Jiye Ai*, Sizhe Feng, Kaori Misuno, Shen Hu*

Division of Oral Biology and Medicine, School of Dentistry, University of California, CHS 63-070 10833 Le Conte Ave Los Angeles, CA 90095-1668 USA

Received: 25 July 2012 Accepted: 30 August 2012 Available Online: 01 September 2012

ABSTRACT

Sjögren's syndrome (SS) is a chronic autoimmune disorder clinically characterized by dry mouth and eyes. The pathogenic mechanism of SS is inadequately understood and a long delay from the start of the symptoms to final diagnosis has been frequently observed. In this paper, we aim to provide an overview about using omics technologies to discover biomarkers for SS diagnosis and understand potential pathways underlying SS pathogenesis. Omics databases relevant to SS such as Sjögren's Syndrome Knowledge Base, Saliva Ontology and SDxMart are also discussed

Keywords: Omics; Proteomics; System Biology; Bioinformatics.

1. Introduction

Sjögren's syndrome (SS), which was first described in 1933 by the Swedish physician Henrik Sjögren [1], is a chronic autoimmune disorder clinically characterized by xerostomia and keratoconjunctivitis sicca. The disease may occur alone as primary SS or present as secondary SS, which is associated with other autoimmune diseases such as rheumatoid arthritis (RA) or systemic lupus erythematosus (SLE). SS has an estimated prevalence of ~ 4 million patients in the US. It primarily affects women, with a ratio of 9:1 over the occurrence in men.

Histologically, SS is characterized by infiltration of exocrine gland tissues with predominantly T lymphocytes, leading to significant reduction of saliva and tear production (dry eyes and mouth). At the molecular level, glandular epithelial cells express high levels of HLA-DR, which has led to the speculation that these cells are presenting antigen (viral antigen or autoantigen) to the invading T cells. Cytokine production follows, with interferon- γ and interleukin-2 being especially important. There is also evidence of B cell activation with autoantibody production and an increase in B cell malignancy. SS patients exhibit a 40-fold increased risk of developing lymphoma than the general population, and the most common forms

are mucosa-associated lymphoid tissue (MALT) lymphomas that remain localized to the salivary glands [2-4].

2. Discovery of saliva biomarkers for SS

Current diagnosis of SS requires an invasive salivary gland tissue biopsy and a long delay from the start of the symptoms to final diagnosis has been frequently observed. There has been increasing interest in developing saliva biomarkers for simple and early diagnosis of the disease. Saliva is the secreted fluid from three pairs of major salivary glands (parotid, submandibular, and sublingual), and multiple minor salivary glands that lie beneath the oral mucosa [5]. It harbors a wide spectrum of analytes - such as proteins, mRNAs, DNAs and metabolites—that may be informative for diagnosis of human diseases. To date, over 1000 distinct proteins in human whole saliva and 1100 proteins from parotid and submandibular/sublingual secretions have been identified [6-10].

Saliva is an attractive medium for disease diagnostics because it is simple to collect and process saliva samples and particularly because salivary testing is non-invasive safe and

*Corresponding authors: Jiye Ai, E-mail Address: jiyelai@gmail.com; Shen Hu, E-mail Address: shenhu@ucla.edu, Phone: number: (+1) 310-206-8834, Fax number: (+1) 310-794-7109

inexpensive. Saliva diagnostics is especially preferred under circumstances where it is difficult to obtain blood samples or repeated sample collection is needed for monitoring a disease. The biologic fluid has been used for the survey of general health and for the diagnosis of diseases in humans, such as human immunodeficiency virus, periodontal diseases, and autoimmune diseases [5, 11-14]. In patients with SS, signature biomarkers from the affected salivary glands may be shed into the lumen and secreted with saliva, which can be identified using modern omics technology. This notion - combined with the inherent advantages of saliva testing - created the emerging interest of developing saliva biomarkers for SS diagnosis, and triggered multiple studies on globally searching for saliva biomarkers of SS.

Proteomics is a powerful approach for global study of the structure and function of all proteins expressed in a biological system. Mapping proteomes from disease tissues and body fluids, has been used to identify new disease biomarkers for clinical and diagnostic applications. By using surface-enhanced laser desorption/ionization time-of-flight mass spec-

trometry and two-dimensional difference gel electrophoresis, Ryu et al. profiled proteins in parotid saliva from primary SS and revealed multiple candidate protein biomarkers for SS, including β -2-microglobulin, lactoferrin, immunoglobulin kappa (κ) light chain, polymeric Ig receptor, lysozyme C and cystatin C in all stages of SS. The study suggested that the salivary proteomic profile of SS is a mixture of increased inflammatory proteins and decreased acinar proteins when compared with non-SS saliva [15]. Hu et al. found that 16 whole saliva (WS) proteins were down-regulated and 25 WS proteins were up-regulated in patients with primary SS compared with matched healthy control subjects. These proteins reflected the damage of glandular cells and inflammation of the oral cavity system in SS patients. In addition, microarray analysis revealed that 27 mRNA in saliva samples were significantly up-regulated in the primary SS patients, most of which are interferon-inducible or related to lymphocytic infiltration and antigen presentation known to be involved in the pathogenesis of primary SS [5]. A panel of these biomarkers has been successfully validated in independent patient population

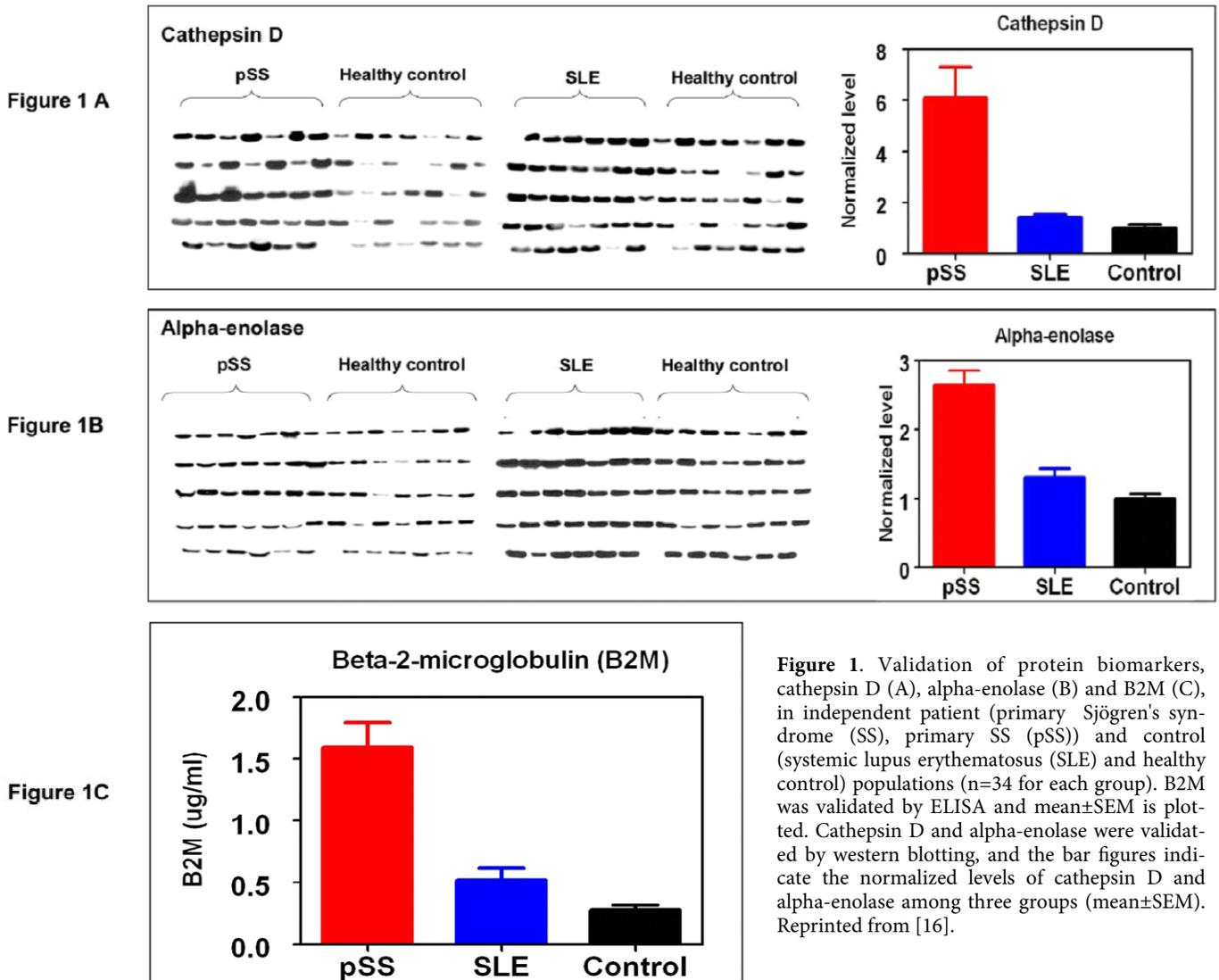


Figure 1. Validation of protein biomarkers, cathepsin D (A), alpha-enolase (B) and B2M (C), in independent patient (primary Sjögren's syndrome (SS), primary SS (pSS)) and control (systemic lupus erythematosus (SLE) and healthy control) populations (n=34 for each group). B2M was validated by ELISA and mean±SEM is plotted. Cathepsin D and alpha-enolase were validated by western blotting, and the bar figures indicate the normalized levels of cathepsin D and alpha-enolase among three groups (mean±SEM). Reprinted from [16].

suggesting that they can discriminate primary SS from both SLE and healthy controls (Figure 1). If further validated in patients with primary SS and those with sicca symptoms but no autoimmune disease, these biomarkers may lead to a simple yet highly discriminatory clinical tool for diagnosis of primary SS [16].

Rigante et al. performed the proteomic analysis of the salivary peptide complex in SS patients' salivary fluid near diagnosis and after 6 months of pharmacological therapy. The analysis revealed that clinical and functional changes of the salivary glands driven by non-steroidal antiinflammatory drugs might be reflected in different proteomic patterns of the salivary fluid [17]. Using a proteomic approach, Giustil et al. characterized the WS proteins of primary SS patients and revealed a set of differentially expressed proteins in patients with primary SS, which are related to acute and chronic inflammation while some others were involved in oxidative stress injury. These findings are in line with the systemic immuno-inflammatory aspects of primary SS and open the possibility for a systematic search of diagnostic biomarkers and targets for therapeutic intervention in primary SS [18]. Peluso et al. investigated the effect of pilocarpine on the salivary peptide and protein profile in patients with primary SS and found that pilocarpine partially restored the levels and numbers of identifiable proteins in saliva from patients with primary SS. Higher levels of alpha-defensin1 and the presence of beta-defensin 2 in the saliva of patients with primary SS could be markers of oral inflammation in this patient group [19].

Using immune-response protoarrays to profile saliva autoantibodies from patients with primary SS or SLE and healthy control subjects, Hu et al. identified salivary autoantibody biomarkers that are highly specific to primary SS. A panel of

24 autoantibody biomarkers was found to be significantly over-produced in primary SS patients compared to both SLE patients and healthy individuals (Figure 2). Four saliva autoantibody biomarkers, anti-transglutaminase, anti-histone, anti-SSA, and anti-SSB, were further tested in independent primary SS (n=34), SLE (n=34), and healthy control (n=34) subjects and all were successfully validated with ELISA. This study has demonstrated the potential of a high-throughput protein microarray approach for the discovery of autoantibody biomarkers. The validated saliva autoantibody biomarkers could well discriminate patients with primary SS from both SLE patients and healthy individuals [20].

3. Systems biology analysis of Sjogren's syndrome

Hu et al. conducted a systems biology analysis of parotid gland tissues obtained from patients with primary SS, patients with primary SS/MALT lymphoma, and subjects without primary SS (non-primary SS controls) (Figure 3). The tissue samples were assessed concurrently by gene-expression microarray profiling and proteomics analysis followed by weighted gene-coexpression network analysis (WGCNA) to identify activated pathways and target genes in respective disease phenotypes. WGCNA is a systems biologic analysis method that has been successfully used to identify disease pathways and their key constituents. It basically identifies gene co-expression modules based on unsupervised clustering of microarray data and explore both gene significance (differential expression) and connectivity for each gene. Gene-coexpression modules related to primary SS or primary SS/MALT lymphoma were significantly enriched with genes known to be involved in the immune/defense response, apoptosis, cell signaling, gene regulation, and oxi-

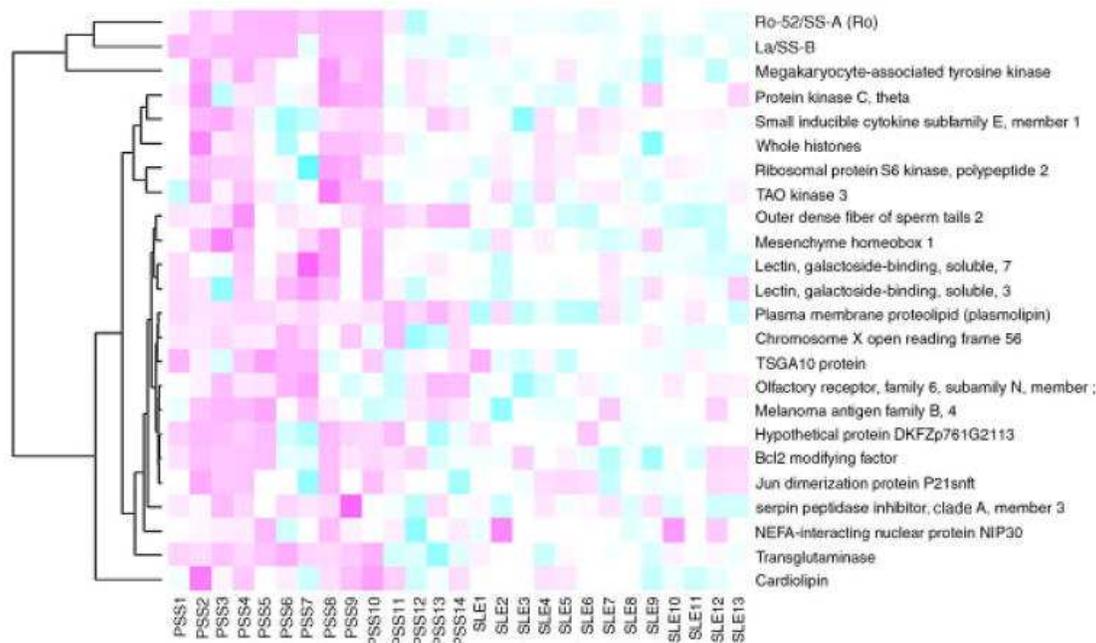


Figure 2. A heatmap of 24 saliva autoantibodies between primary Sjögren's syndrome (pSS) and SLE groups based on the protein microarray profiling. Reprinted from [20].

ductive stress. Detailed functional pathway analyses indicated that primary SS-associated modules were enriched with genes involved in proteasome degradation, apoptosis, signal peptides of the class I major histocompatibility complex (MHC), complement activation, cell growth and death, and integrin-mediated cell adhesion, while primary SS/MALT lymphoma-associated modules were enriched with genes involved in translation, ribosome biogenesis and assembly, proteasome degradation, class I MHC signal peptides, the G13 signaling pathway, complement activation, and integrin-mediated cell adhesion. Combined analyses of gene expression and proteomics data implicated 6 highly connected "hub" genes for distinguishing primary SS from non-primary SS, and 8 hub genes for distinguishing primary SS/MALT lymphoma from primary SS. The identified gene modules/pathways provide further insights into the molecular mechanisms of primary SS and primary SS/MALT lymphoma whereas the identified disease-hub genes represent promising targets for therapeutic intervention, diagnosis, and prognosis [21].

4. Bioinformatics

State-of-the-art omics technologies, including proteomics

and transcriptomics, are being implemented widely in studies of human disease including SS. Bioinformatics approaches such as mining the data from multiple omics studies can provide deeper insights into the entire systems than can be obtained from any single omics study. This section introduces several bioinformatics infrastructures relevant to SS.

The Sjögren's syndrome knowledge base [22] (SSKB <http://sskb.umn.edu>) is a database that collects and organizes gene and protein expression data from the existing literature for comparative analysis with future gene expression and proteomic studies of SS. The SSKB is generated from PubMed using text mining of over 7,700 abstracts and listing approximately 500 potential genes/proteins. The SSKB can be used for literature reviews and literature-based validation of identified genes, functional gene enrichment studies, protein-protein interaction networks and other bioinformatics analyses.

Saliva Ontology (SALO) is a consensus-based controlled vocabulary of terms and relations dedicated to the omics domain and to saliva-related diagnostics. SALO is tested specifically in light of its capacity to meet the ontology needs for managing data derived from research on the use of a salivary marker for SS. SS-relevant portion of the ontology is to be validated through the work on annotation of representa-

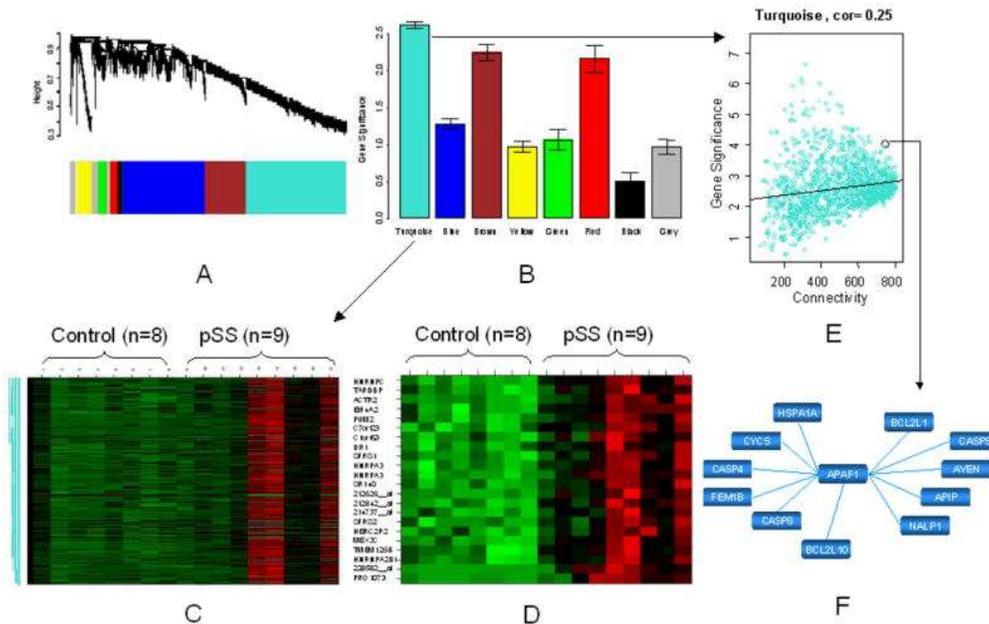


Figure 3. Module and hub gene detection based on gene expression analysis of primary SS (pSS) patients and controls. (A) Hierarchical cluster tree used for module detection. Modules correspond to branches of the tree and are assigned the same color as indicated by the color-band underneath the tree. Grey denotes genes outside proper modules; (B) Disease-related gene co-expression modules identified by WGCNA. Module significance is defined as average gene significance and the gene significance measure is defined as the absolute value of the Student T statistic for differential expression between pSS and control groups. Note that the Turquoise, Brown, and Red modules are comprised of highly differentially expressed genes (average absolute T-test > 2). (C) Heatmap of the gene expression data of the Turquoise module genes (rows) versus array samples (y-axis). The columns on the left hand side correspond to controls and those on the right hand side to pSS subjects. Note the Turquoise module genes tend to be highly over expressed in pSS patients. (D) Heatmap of the 22 most significant genes from the Turquoise module which can well segregate pSS and control groups. (E) Scatter plot of gene significance versus intramodular connectivity in the Turquoise module. While highly connected intramodular hub genes tend to be differentially expressed, connectivity (module membership) and gene significance are complementary gene screening variables. (F) One particularly promising hub gene in the Turquoise module, apoptotic peptidase activating factor 1 (APAF1), encodes a cytoplasmic protein that initiates apoptosis. Reprinted from [21].

tive research literature on the SSKB as well as our research on the salivary protein biomarkers[23].

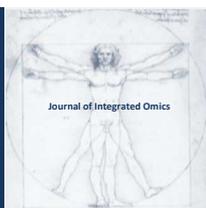
SDxMart is a BioMart[24] data portal that hosts saliva omics data and offers access to the data by using the BioMart interface and querying environment. The SDxMart is designed to provide a variety of queries to facilitate saliva biomarker discovery including complex queries that integrate omics, clinical and functional information. The SDxMart holds data from projects of oral diseases and systemic diseases including SS. The types of omics datasets are proteomics, transcriptomics, microRNA and metabolomics[25].

Acknowledgements

The work was supported by the Sjögren's Syndrome Foundation.

References

- H. Sjogren. Zur Kenntnis der Keratoconjunctivitis sicca. *Acta Ophthalmol.* 1933;11:Suppl 2: 1-151.
- R. I. Fox. Sjögren's syndrome. *Lancet.* 2005 2005/7/29;366 (9482):321-31.
- R. I. Fox, H. Kang, D. Ando, J. Abrams, E. Pisa. Cytokine mRNA expression in salivary gland biopsies of Sjogren's syndrome. *J Immunol.* 1994;152:5532-9.
- A. Hansen, P. E. Lipsky, T. Dorner. B-cell lymphoproliferation in chronic inflammatory rheumatic diseases. *Nat Clin Pract Rheum.* 2007;3(10):561-9.
- S. Hu, J. Wang, J. Meijer, S. Jeong, Y. Xie, T. Yu, H. Zhou, S. Henry, A. Vissink, J. Pijpe, C. Kallenberg, D. Elashoff, J. A. Loo, D. T. Wong. Salivary proteomic and genomic biomarkers for primary Sjogren's syndrome. *Arthritis Rheum.* 2007 Nov;56 (11):3588-600.
- S. Hu, Y. Xie, P. Ramachandran, R. R. Ogorzalek Loo, Y. Li, J. A. Loo, D. T. Wong. Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. *Proteomics.* 2005 Apr;5(6):1714-28.
- S. Hu, Y. Li, J. Wang, Y. Xie, K. Tjon, L. Wolinsky, R. R. Loo, J. A. Loo, D. T. Wong. Human saliva proteome and transcriptome. *J Dent Res.* 2006 Dec;85(12):1129-33.
- S. Hu, J. A. Loo, D. T. Wong. Human saliva proteome analysis. *Ann N Y Acad Sci.* 2007 Mar;1098:323-9.
- H. Xie, N. L. Rhodus, R. J. Griffin, J. V. Carlis, T. J. Griffin. A catalogue of human saliva proteins identified by free flow electrophoresis-based peptide separation and tandem mass spectrometry. *Mol Cell Proteomics.* 2005 Nov;4(11):1826-30.
- T. Guo, P. A. Rudnick, W. Wang, C. S. Lee, D. L. Devoe, B. M. Balgley. Characterization of the human salivary proteome by capillary isoelectric focusing/nanoreversed-phase liquid chromatography coupled with ESI-tandem MS. *J Proteome Res.* 2006 Jun;5(6):1469-78.
- D. Malamud. Oral diagnostic testing for detecting human immunodeficiency virus-1 antibodies: a technology whose time has come. *Am J Med.* 1997 Apr 1;102(4A):9-14.
- E. Kaufman, I. B. Lamster. Analysis of saliva for periodontal diagnosis--a review. *J Clin Periodontol.* 2000 Jul;27(7):453-65.
- W. W. Kalk, A. Vissink, B. Stegenga, H. Bootsma, A. V. Nieuw Amerongen, C. G. Kallenberg. Sialometry and sialochemistry: a non-invasive approach for diagnosing Sjogren's syndrome. *Ann Rheum Dis.* 2002 Feb;61(2):137-44.
- J. Pijpe, W. W. Kalk, H. Bootsma, F. K. Spijkervet, C. G. Kallenberg, A. Vissink. Progression of salivary gland dysfunction in patients with Sjogren's syndrome. *Ann Rheum Dis.* 2007 Jan;66(1):107-12.
- O. H. Ryu, J. C. Atkinson, G. T. Hoehn, G. G. Illei, T. C. Hart. Identification of parotid salivary biomarkers in Sjogren's syndrome by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry and two-dimensional difference gel electrophoresis. *Rheumatology (Oxford).* 2006 Sep;45(9):1077-86.
- S. Hu, K. Gao, R. Pollard, M. Arellano-Garcia, H. Zhou, L. Zhang, D. Elashoff, C. G. Kallenberg, A. Vissink, D. T. Wong. Preclinical validation of salivary biomarkers for primary Sjogren's syndrome. *Arthritis Care Res (Hoboken).* 2010 Nov;62(11):1633-8.
- D. Rigante, R. Inzitari, M. Carone, C. Fanali, A. Stabile, T. Cabras, E. Capoluongo, S. Gaspari, G. Barone, M. Castagnola. Correspondence between clinical improvement and proteomic changes of the salivary peptide complex in a child with primary Sjogren syndrome. *Rheumatol Int.* 2008 Jun;28(8):801-6.
- L. Giusti, C. Baldini, L. Bazzichi, F. Ciregia, I. Tonazzini, G. Mascia, G. Giannaccini, S. Bombardieri, A. Lucacchini. Proteome analysis of whole saliva: a new tool for rheumatic diseases--the example of Sjogren's syndrome. *Proteomics.* 2007 May;7(10):1634-43.
- G. Peluso, M. De Santis, R. Inzitari, C. Fanali, T. Cabras, I. Messana, M. Castagnola, G. F. Ferraccioli. Proteomic study of salivary peptides and proteins in patients with Sjogren's syndrome before and after pilocarpine treatment. *Arthritis Rheum.* 2007 Jul;56(7):2216-22.
- S. Hu, A. Vissink, M. Arellano, C. Roozendaal, H. Zhou, C. G. Kallenberg, D. T. Wong. Identification of autoantibody biomarkers for primary Sjogren's syndrome using protein microarrays. *Proteomics.* 2011 Apr;11(8):1499-507.
- S. Hu, M. Zhou, J. Jiang, J. Wang, D. Elashoff, S. Gorr, S. A. Michie, F. K. Spijkervet, H. Bootsma, C. G. Kallenberg, A. Vissink, S. Horvath, D. T. Wong. Systems biology analysis of Sjogren's syndrome and mucosa-associated lymphoid tissue lymphoma in parotid glands. *Arthritis Rheum.* 2009 Jan;60 (1):81-92.
- S. U. Gorr, T. J. Wennblom, S. Horvath, D. T. Wong, S. A. Michie. Text-mining applied to autoimmune disease research: the Sjogren inverted question marks syndrome knowledge base. *BMC Musculoskelet Disord.* 2012 Jul 3;13(1):119.
- J. Ai, B. Smith, D. T. Wong. Saliva Ontology: an ontology-based framework for a Salivaomics Knowledge Base. *BMC Bioinformatics.* 2010;11:302.
- J. M. Guberman, J. Ai, O. Arnaiz, J. Baran, A. Blake, et. al.. BioMart Central Portal: an open database network for the biological community. *Database (Oxford).* 2011;2011:bar041.
- J. Y. Ai, B. Smith, D. T. Wong. Bioinformatics advances in saliva diagnostics. *Int J Oral Sci.* 2012 Jun 15;4.



REVIEW ARTICLE | DOI: 10.5584/jiomics.v2i2.105

A survey on coronary heart disease related signal pathways, drug targets and pharmacological interventions

Peng Jiang¹, Runhui Liu^{1,*}, Weidong Zhang^{1,2,**}

¹School of Pharmacy, Second Military Medical University, Shanghai, P.R. China; ²School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200030, P.R. China; ³Shanghai Hutchison Pharmaceuticals Company, Shanghai 200331, P.R. China.

Received: 19 August 2012 Accepted: 02 October 2012 Available Online: 01 November 2012

ABSTRACT

Coronary heart disease (CHD), the most common form of cardiovascular disease, is a chronic, multifactorial disease. With significant advances in our understanding of the pathophysiological process of CHD in recent years, more and more drug targets have been identified and adopted in drug discovery for CHD. In this review, a comprehensive perspective of pathological and pharmacological development of CHD was introduced through searching in multiple bibliographic sources. CHD related signal pathways (including 409 proteins), drug targets (including 101 proteins) and pharmacological interventions were summarized and visualized. The knowledge of these signal pathways and drug targets may facilitate new drug discovery and medicine intervention for CHD in the future.

1. Introduction

Cardiovascular disease (CVD) is the most common cause of death worldwide. It was estimated that around 23.6 million people will die from CVDs by 2030 and approximately half of these occurrences are directly related to coronary heart disease (CHD)[1,2]. CHD is the consequence of atherosclerosis with accumulation of atheromatous plaques within walls of coronary arteries that supply myocardium with oxygen and nutrients. The progress of CHD is characterized by its chronicity. It is particularly insidious in initial stages. Several risk factors such as cigarette smoking, hypercholesterolemia, hypertension, hyperglycemia and work stress have been demonstrated to be closely related to the processing of CHD. Acute myocardial infarction, arrhythmia, angina pectoris and heart failure are major clinical manifestation of CHD.

In the formation of atherosclerosis and CHD, endothelial cells (ECs), vascular smooth muscle cells (VSMCs), macrophages, T leukomonocytes, monocytes, mast cells, dendritic

cells, platelets and cardiocytes interact with each other to damage normal functions of coronary artery and heart muscle [3-10]. Multiple intracellular and extracellular signal pathways containing hundreds of proteins participate in these interactions. A total of 413 CHD associated proteins were listed in Supplementary Table 2. Some of these proteins, such as 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase), angiotensin-conversion enzyme (ACE), and angiotensin (Ang II) receptor, have been identified as drug targets. Drugs designed according to these targets have been playing important roles in the treatment of CHD.

In this review, we summarized CHD related signal pathways, drug targets and pharmacological interventions (including combination therapy) in clinic through searching in multiple bibliographic sources. By doing so, a comprehensive knowledge of ligands and targets relevant to CHD can be shown and further facilitate the drug discovery.

*Corresponding author: Associate Professor Runhui Liu, School of Pharmacy, Second Military Medical University, No. 325 Guohe Road, Shanghai, 200433, P.R. China, E-mail Address: lyliurh@126.com, Fax Number: +86-21-81871245; Professor Weidong Zhang, School of Pharmacy, Second Military Medical University, No. 325 Guohe Road, Shanghai, 200433, P.R. China, E-mail Address: wdzhangy@hotmail.com, Fax Number: +86-21-81871244

2. Cells, pathways and proteins related to CHD

CHD is a chronic disease in which blood flow is obstructed through coronary arteries that supply heart with oxygen-rich blood. As shown in Fig. 1, this obstruction is caused by atherosclerosis, whose development is a lifelong process. According to “response to injury” hypothesis of atherosclerosis, endothelial dysfunction is the first step in atherosclerosis [11]. Subsequently, monocytes gather at the site of the injury and in turn provoke an inflammatory immune response that causes further damage to arterial wall. Over time, low density lipoprotein (LDL) penetrates into arterial wall and is modified by oxidasis, then combining with the monocytes derived macrophages to form foam cell. Simultaneously, T lymphocytes, dendritic cells, mast cells and platelets can enter the intima of arterial wall through injured endothelial layer and induce the releasing of cytokines. With the participation of these cytokines such as matrix metalloproteinases (MMPs) and platelet-derived growth factors (PDGF), VSMCs can proliferate and migrate through intermediate lesion via the degradation of extracellular matrix components, thus promoting the formation of atherosclerosis plaque. Once plaque ruptures, some pieces of the plaque can

travel through arteries until causing a blockage. Then, myocardial infarction happens and leads to dysfunction of cardiocytes. Finally, injured vascular and heart muscle lead to the occurrence of CHD.

In this section, we mainly introduce the functions of related cells in the formation of CHD, including ECs, VSMCs, T lymphocytes, dendritic cells, monocytes, macrophage, mast cells, platelets and cardiocytes.

2.1 ECs

ECs are inert barriers that separate flowing blood from underlying tissue. They play important roles in regulating hemostasis, cellular and nutrient trafficking. Under normal conditions, ECs are well aligned, tight junction with very low rates of death and permeable to some macromolecules such as LDL. ECs can produce numerous vasoactive factors, such as nitric oxide (NO), prostaglandin, endothelin (ET-1) and Ang II. Vascular dysfunction due to endothelial cell injury alters these normal homeostatic properties. For example, the synthesis of NO is reduced, whilst the expression of vasoconstrictive molecules such as ET-1, is up-regulated [12, 13].

Hyperlipidemia, hypertension, hyperglycemia and smok-

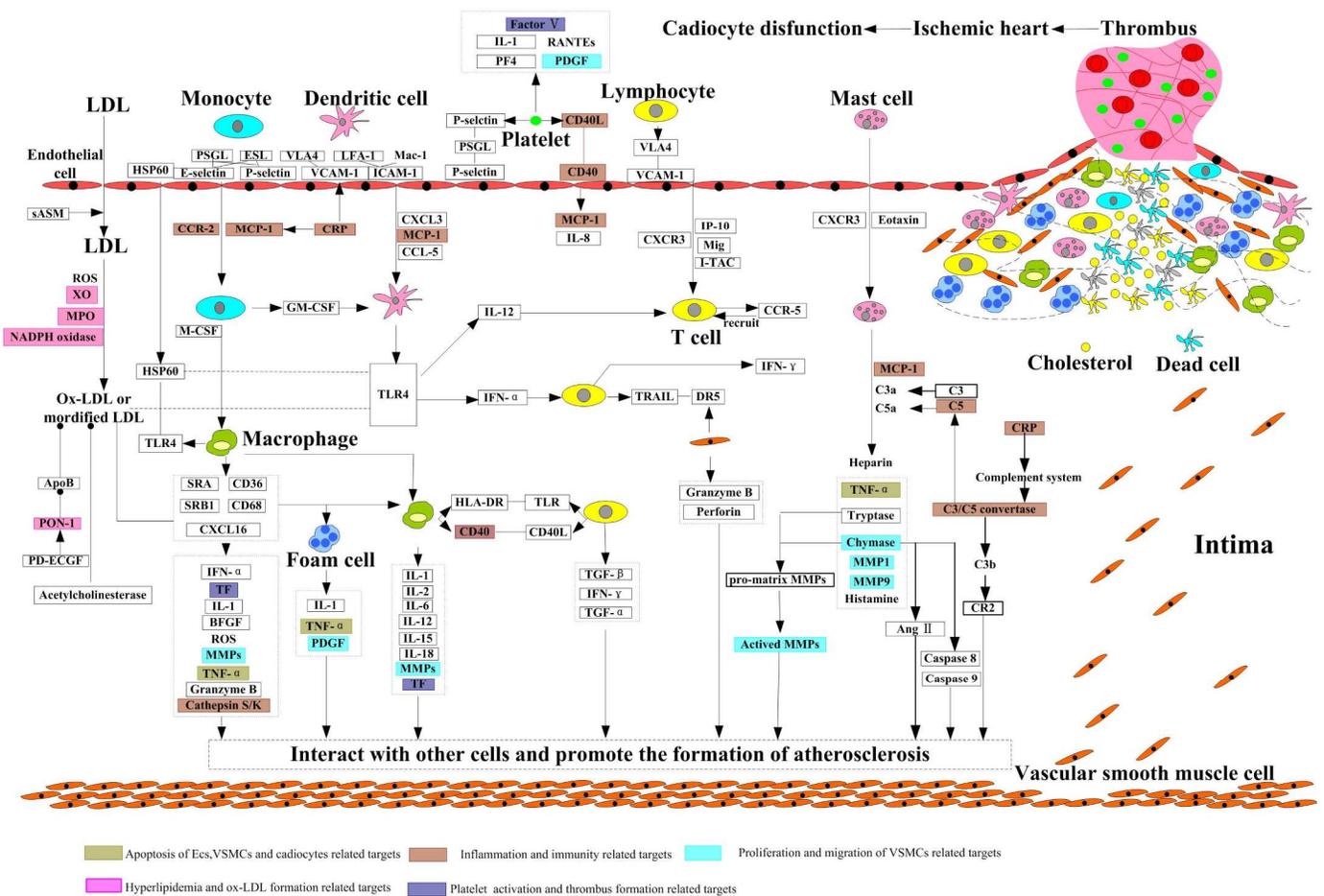


Fig. 1 Interaction of different cells in the formation of ischemic heart. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

ing are main causes for ECs injury. As the increased permeability of ECs induced by injury, atherogenic matter LDL transmigrates through the endothelial layer and transforms into ox-LDL[14]. As shown in Fig. 2, once ox-LDL enters the intima, oxidized low-density lipoprotein receptor 1 (LOX) can be activated, thus inducing NF-Kb activation and leading to the up-regulation of adhesion molecules such as E-selectin, P-selectin, platelet-endothelial cell adhesion molecule (PECAM), intercellular adhesion molecule (ICAM) and chemotatic factors such as C-C motif chemokine 2 (MCP-1) and macrophage colony-stimulating factor 1 (M-CSF). In fact, except ox-LDL, HSP 60 and cytokines such as arachidonic acid (AA), protein geranylgeranyltransferase type-1 (GGTase), advanced glycation endproducts (AGEs), immunoreactive fibronectin-γ (IFN-γ) and interleukin-1 (IL-1) can also promote ECs to secrete adhesion molecules and chemotatic factors. With the help of these adhesion mole-

cles and chemotatic factors, the atherosclerotic cells like monocytes, lymphocytes, dendritic cells, mast cells and platelets begin to accumulate and enter intima, thus increasing the likelihood of plaque formation [14-16].

Apoptosis of ECs is an important event in ECs injury. As Fig. 2 shows, ox-LDL not only promotes the accumulation of atherogenic cells, but also participates in the apoptosis of ECs by activating caspase pathway. Oxidative stress and tumor necrosis factor-α(TNF-α) are also important factors leading to the apoptosis of ECs through regulating caspase pathway.

2.2 VSMCs

VSMCs are essential for vascular contraction and relaxation. They alter luminal diameter and enable blood vessels to maintain an appropriate blood pressure. The increased vas-

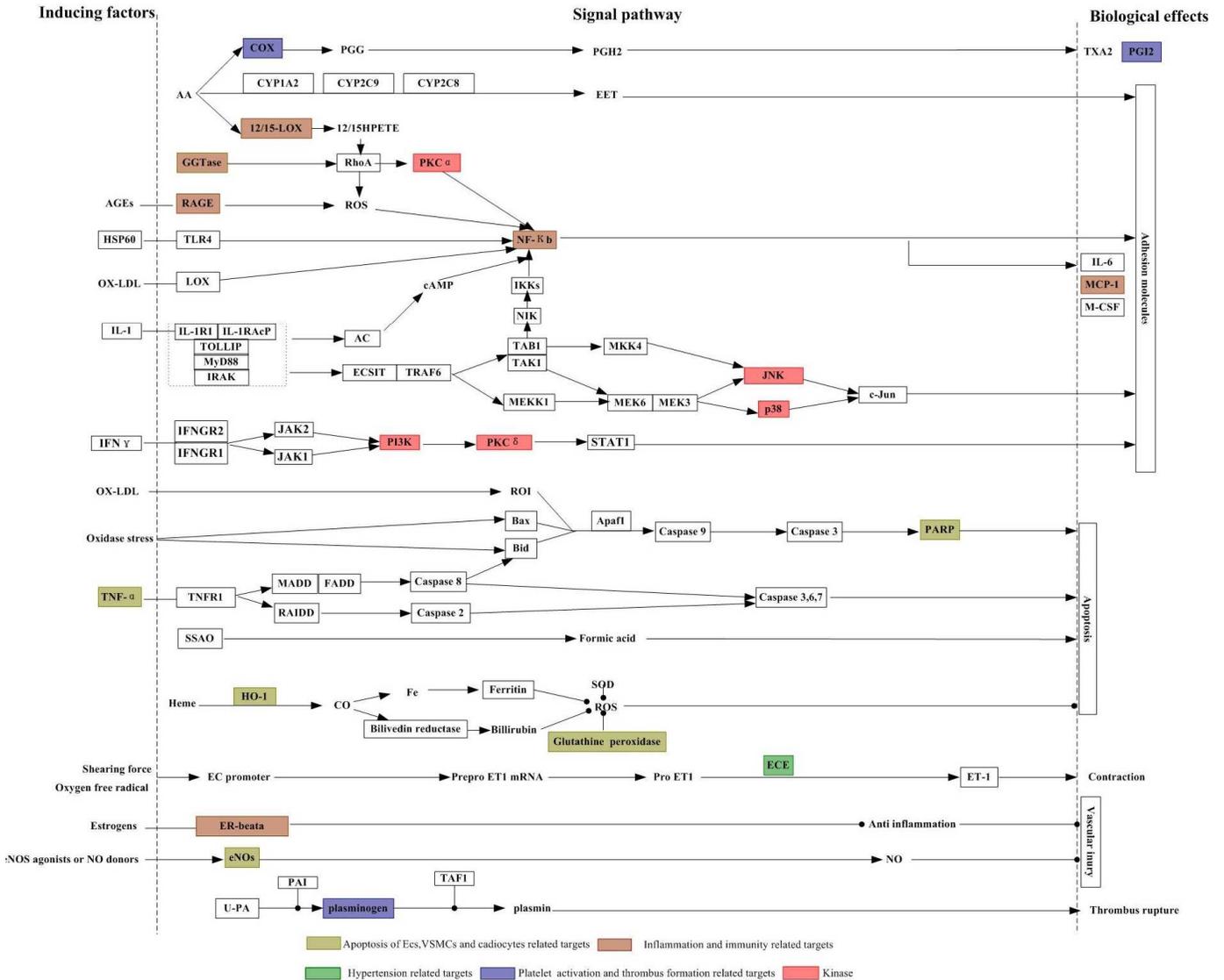


Fig. 2 The intracellular signal pathway of endothelial cell in CHD. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

cular contractility, migration, proliferation and apoptosis of VSMCs are all important factors for the formation of atherosclerotic plaque.

(1) Contractility: The activation of renin-angiotensin system (RAS) is an important event for contractility of VSMCs. Once RAS is activated, the level of renin in blood is evaluated, thus promoting the production of Ang II. By combining with Ang II receptors (AgtR1 and AgtR2), the contraction of vascular increase significantly (shown in Fig. 3) [17]. ET-1, the strongest vasoconstrictor peptide secreted by impaired ECs, also can stimulate contraction of VSMCs [18]. In addition, as shown in Fig. 3, 5-hydroxytryptamine (5-HT) and AA metabolites such as hydroxy-eicosatetraenoic acids (HETEs) also make a contribution to the contraction of VSMCs.

(2) Migration and proliferation: As shown in Fig. 3, PDGF secreted by foam cells and activated platelets is the most important factor for migration and proliferation of VSMCs [19]. MMPs which can digest the extracellular matrix (ECM) in the intima also supply convenient condition for migration of VSMCs. In addition, urokinase-type plasminogen activator (U-PA) and leukotriene B4 receptor 1 (BLT1) also can aggravate the migration of VSMCs through activating focal adhesion kinase 1 (FAK) and up-regulating the production of MMPs (shown in Fig. 3)[20, 21].

In the aspect of proliferation of VSMCs, except PDGF, Ang II and ET-1 are also important factors for the proliferation for VSMCs (shown in Fig. 3). In recent years, some studies demonstrated that urotensin II can also inducing the proliferation of VSMCs [22, 23].

(3) Apoptosis: Proliferation and apoptosis of SMCs are coincidence events which not only contribute to vessel remodeling but also lead to destabilization of fibrous cap in arteriosclerotic lesions[24, 25]. Various stimuli, including oxidized lipoproteins, altered hemodynamic stress and free radicals, can precipitate macrophage and T lymphocytes to secrete apoptosis factors. TNF- α and IFN- γ are typical factors which contribute to the apoptosis of VSMCs and ECs [24] (shown in Fig. 2). T lymphocytes can also induce apoptosis of VSMCs via releasing perforin and granzyme B (shown in Fig. 1). In addition, Fas mediated apoptosis of VSMCs could be another important pathway as reported[26] (shown in Fig. 3).

2.3 Monocytes and macrophages

Monocytes can enter intima and differentiate into macrophages in the formation of atherogenesis. As shown in Fig. 1, with the continuing expression of adhesion molecules on injured ECs, monocytes are initially attracted to lesion-prone sites. The initial adhesion involves selectins, which mediate a rolling interaction, and is followed by firmer attachment by means of integrins. Adherent monocytes migrate into intima with the help of chemoattractant molecules MCP-1[4, 27]. After migrating into the subendothelial space, monocytes differentiate into activated macrophages via the existence of

M-CSF[28]. Once differentiation is finished, scavenger receptors such as scavenger receptor class A (SRA) and platelet glycoprotein 4 (CD36) on surface of macrophages will phagocytose ox-LDL or other modified LDL, thus inducing the formation of foam cells and fatty streak in atherosclerosis.

Macrophages are considered to be major inflammatory mediators during atherosclerosis progression. As shown in Fig. 1 and Fig. 4, through combination of CD40 ligand with its receptor, macrophages can express amount of chemokines and cytokines such as MCP-1, RANTES, IL-1, IFN- γ , MMPs, TNF- α and tissue factor[29]. Once macrophages develop into foam cells, they also secrete amount of cytokines. These factors continually augment inflammatory reaction in vessels and promote the development of atherosclerosis[5, 15] (Fig. 1). Except playing an important role in inflammation, macrophages also mediate the immunity reaction in atherosclerosis. Macrophages which contain phagotrophic ox-LDL particles act as antigen-presenting cells to T-cells [6] (Fig. 1).

Besides, a number of recent studies have demonstrated that macrophages also play an important role in reverse cholesterol transport [30, 31]. Ox-LDL can be decomposed into cholesterol and oxysterol in macrophages. Cholesterol is further modified into cholesterol ester with the help of acyl-CoA cholesterol acyl transferase (ACAT) and this leads to the formation of foam cells. Oxysterol can activate liver X receptor (LXR) in macrophages, thus promoting cholesterol efflux through apolipoprotein A-1 (ApoA1). By up-regulating ATP-binding cassette transporter 1 (ABCA1) and ApoA1, retinoic acid X activated peroxisome proliferator-activated receptor (PPAR) also participates in reverse transport of cholesterol in macrophages. In addition, nuclear receptor ROR, farnesoid X receptor (FXR) and ADP-ribosylation factor-related protein 1 (ARF1) in macrophages can also regulate the transportation of cholesterol through activating or inhibiting ApoA1.

2.4 T lymphocytes

T lymphocytes are the most important immune cells in atherosclerosis plaque. The trans-endothelial process of lymphocytes is similar to that of monocytes except using different chemoattractants. Known chemoattractants for T lymphocytes include inducible protein-10 (IP-10), monokine induced by IFN- γ (Mig) and IFN-inducible T-cell α -chemoattractant (I-TAC) [15, 32]. These chemokines bind to C-X-C chemokine receptor type 3 (CXCR3) which is expressed by T lymphocytes in atherosclerotic lesion and facilitate the migration of T lymphocytes (Fig.1).

Once residence in the arterial intima, T lymphocytes have chances to interact with antigen-presenting cells (APCs) such as macrophages. More and more studies demonstrated that CD40 ligand (CD40L) and its receptor CD40, which can be expressed in macrophages and T cells, play an important role in antigen presentation and autoimmunity as costimulatory factors [33, 34]. Except macrophages, VSMCs

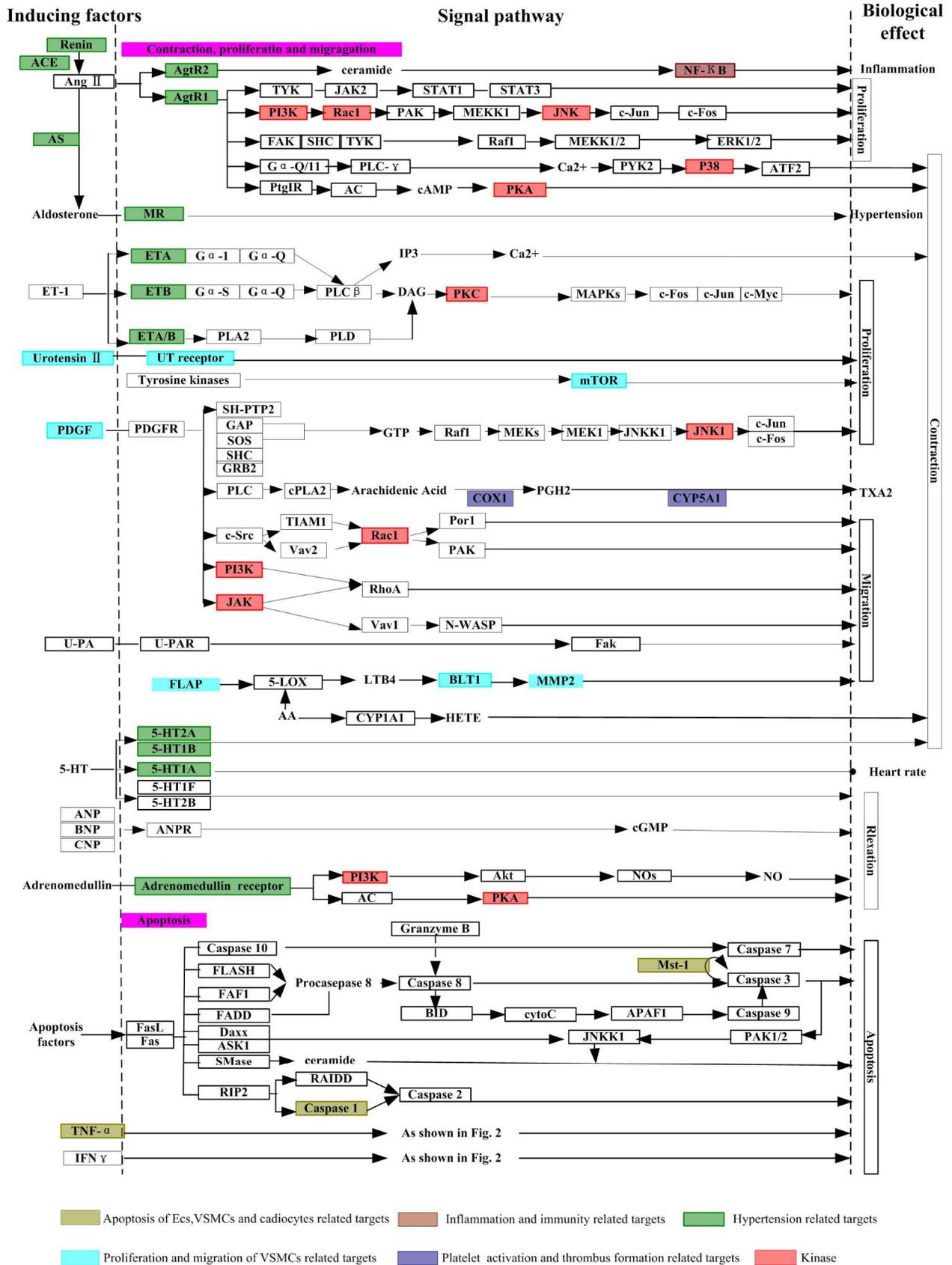


Fig. 3 The intracellular signal pathway of vascular smooth muscle cell in CHD. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

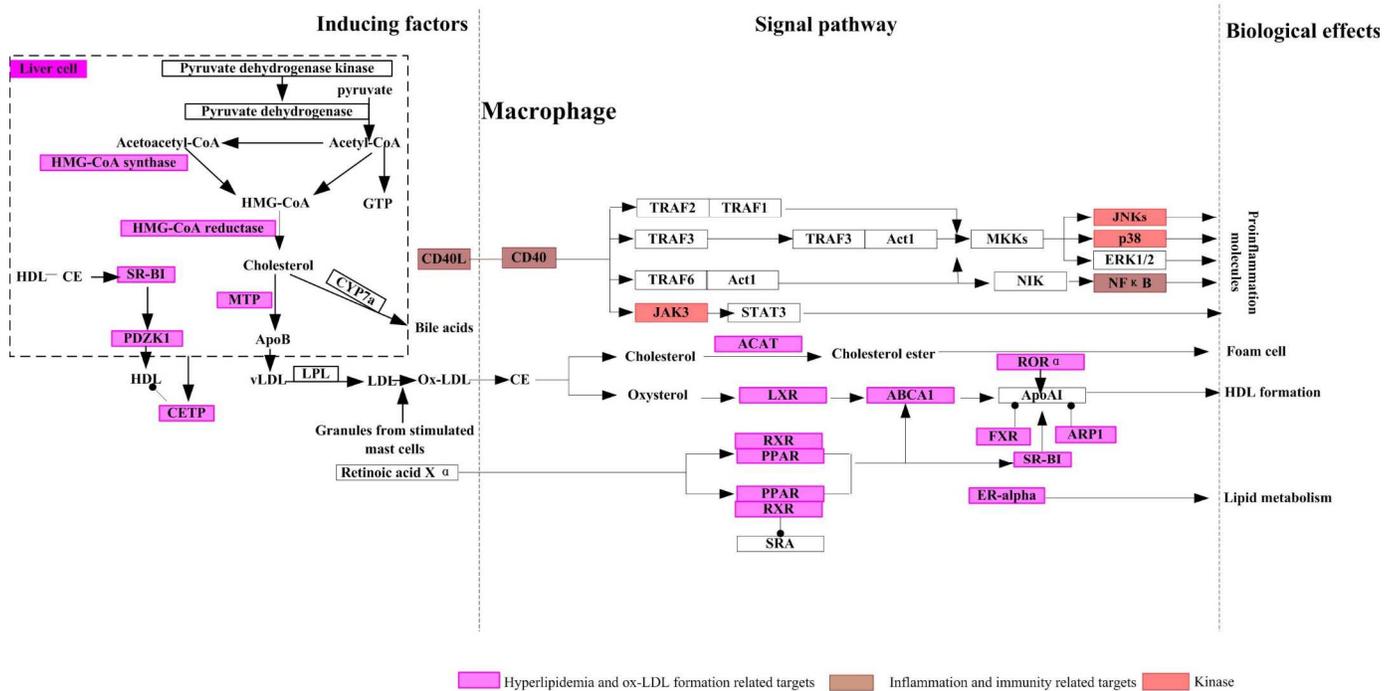


Fig. 4 The intracellular signal pathway of macrophage in CHD. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

are another kind of APCs which can be recognized by T cells. By binding to a HLA class II histocompatibility antigen (DR5) which is expressed by VSMCs, T cells are activated, and then secrete granzyme B and perforin, both of which can induce apoptosis of VSMCs [8, 26](Fig.1). In some studies, oxidized LDL and heat-shock protein 60 (HSP60) also have been identified as antigens for T lymphocytes [35, 36].

As can be seen in Fig. 1, activated T cells also predominates the production of pro-inflammatory cytokines such as IFN- γ , TNF α and β and chemokines like C-C chemokine receptor type 5 (CCR5). These cytokines can augment the inflammatory and immune response [37].

2.5 Mast cells

Mast cells, an inflammatory cell type, contain highly enriched proteases, tryptase and chymase in intracellular granules. They have been found to participate in the inflammatory reaction of atherosclerotic lesions.

As shown in Figure 1, the trans-endothelial migration of mast cells is mediated by a chemoattractant named eotaxin which interacts with the chemokine receptor CXCR3 expressed on the surface of mast cells[32]. After entering intima and then being activated by complement system components such as complement C3a, complement C5a and chemokines such as MCP-1 [38, 39], mast cells become degranulation and release a number of inflammatory mediators including histamine, tryptase, chymase and a variety of cytokines, which can promote vascular inflammation, endothelial dysfunction and foam cell formation[40]. The subendothelial LDL can be modified by binding to the granule rem-

nants from activated mast cells. These binding particles can be phagocytosed by macrophages more likely than LDL alone, thus promoting the formation of foam cells [41] (Fig. 4). A study has demonstrated that granules from stimulated mast cells can also degrade the capability of removing cholesterol from macrophages [42]. Therefore, mast cells not only promote LDL aggregation but also interfere with cholesterol removal, both of which contribute to foam cell formation.

Mast cells can also weaken the fibrous cap in different ways as follows (Fig. 1). (1) Activated mast cells may release MMPs, such as MMP-1 and MMP-9, both of which directly cause matrix degradation [43, 44]. (2) Secreted tryptase and chymase precipitate the pro- MMPs to form active MMPs [40, 45]. (3) Stimulated mast cells can express TNF- α which is able to enhance the apoptosis of VSMCs, macrophages and ECs, and subsequently weaken and rupture the atherosclerotic plaques [46]. (4) Chymase released from mast cells is reported to activate caspase-8 and caspase-9, two key effector molecules in apoptotic cascade[47].

2.6 Dendritic cells

DCs are also antigen-presenting cells with the unique ability to initiate a primary immune response by activating naive T lymphocytes. Though, DCs are typically localized in the subendothelial space as indigenous residents of healthy arteries, circulating DCs in blood can evade into injury sites with the help of chemokines and adhesion molecules in the progress of atherosclerosis[48]. As shown in Fig. 1, adhesion molecules such as P-selectin, E-selectin and VCAM-1 and

chemokines such as C-X-C motif chemokine 3 (CXCL3) and C-C motif chemokine 5 (CCL5) are all potential candidates for recruiting DCs to trans-endothelium[49, 50]. Granulocyte/macrophage colony-stimulating factor (GM-CSF), which may facilitate the transformation of trans-endothelial monocytes into DCs, is another factor to increase the number of DCs in atherosclerotic area [51].

Subendothelium DCs express toll like receptors (TLRs) to recognize dangerous signals and phagocytize antigens such as oxidized LDL and heat-shock proteins. Once finishing the process of phagocytosis, DCs become activated and produce vast amounts of cytokines such as IFN- α and interleukin-12 (IL-12). IFN- α can up-regulate the expression of the proapoptotic protein TRAIL on T lymphocytes thereby multiplying their ability to kill plaque resident cells [8, 52]. IL-12 can up-regulate the expression of chemokine receptor CCR5 on T lymphocytes, which in turn leads to the accumulation of T lymphocytes in the atherosclerotic plaque [53].

2.7 Platelets

Platelets play an important role in hemostatic process and thrombus formation, as well as in the inflammation reaction in atherosclerotic plaque.

After the injury of ECs, the endothelial cells' barrier is lost. Extracellular matrix leaks from intima and triggers the formation of a hemostatic thrombus. In this event, platelets participate in three successive and closely integrated biological process, i.e., adhesion, activation and aggregation [54] (Fig. 5).

(1) At impaired vascular lesions, extracellular matrix components like von Willebrand factor (vWF) and collagen are exposed to the blood. Platelets can detect these changes in blood circulation and adhere to collagen via the membrane

adhesion receptor GPVI (platelet glycoprotein VI) which lead to further combination of vWF with integrin receptors GpIb-IX-V (platelet glycoprotein IX-V) and collagen with $\alpha 2\beta 1$ (integrin alpha-2) in platelets, thus resulting in the fast adhesion [55-58].

(2) Once adhesive receptors combine with their ligands on extracellular matrix components, the activation of platelets will start. This process can be further strengthened by thrombin, adenosine diphosphate (ADP), epinephrine and thromboxane A2 (TXA2), all of which are synthesized by stimulated platelets [54, 59-61].

(3) Aggregation is mediated by adhesive substrates bounding to membranes of activated platelets. Among adhesive substrates, activated GPIIb/IIIa (integrin beta-3/integrin alpha-IIb) is one of the most important factors contributing to the stable adhesion and recruiting more platelets [54, 62].

Platelets are also mediators of inflammation involved in the development of atherosclerosis. As shown in Fig. 1, (1) activated platelets can release P-selectin, which promote monocytes recruitment via platelet-monocyte interactions and deliver platelets' proinflammatory factors to monocytes; (2) activated platelet surface also express CD40L, which further binds to CD40 on the surface of ECs to up-regulate the expression of adhesion molecules and chemokines in ECs [63]; (3) Activated platelets can release amount of pro-inflammatory cytokines, chemokines, growth factors and blood coagulation factors, thus promoting leukocytes recruitment and proliferation of VSMCs 17 (Fig. 2).

2.8 Cadiocytes

Once obstruction happens in coronary vessels, myocardial ischemia takes place and disturbs the normal function of cadiocytes including dysfunction of myocardial energy me-

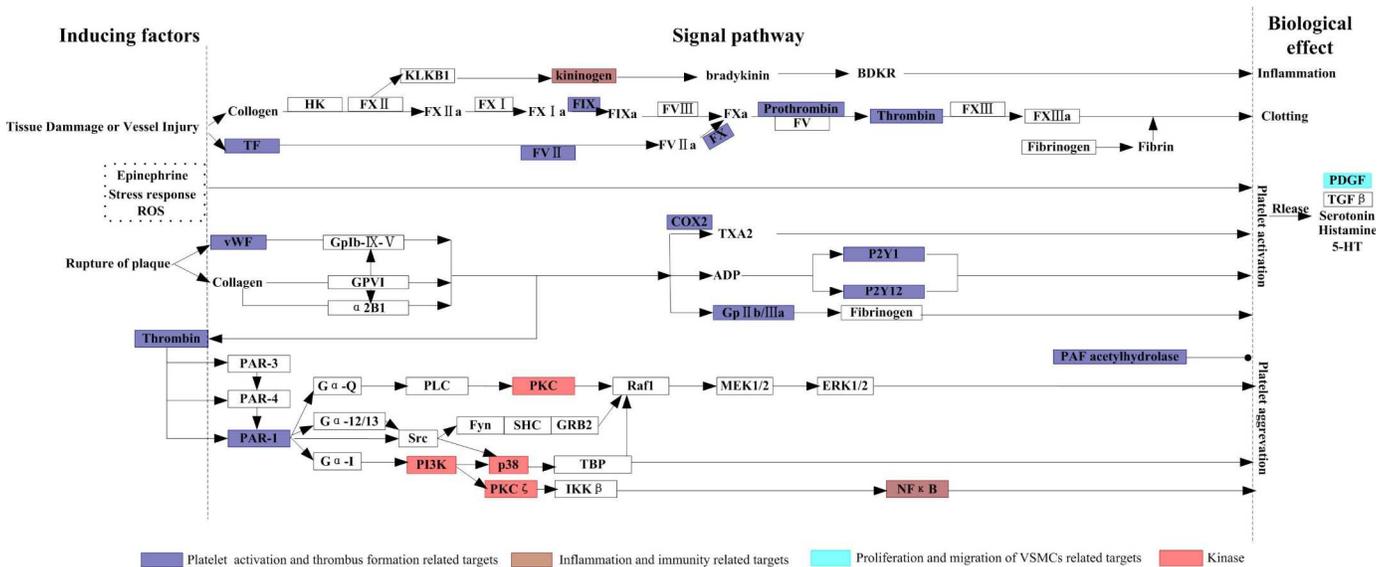


Fig. 5 The intracellular signal pathway of platelet in CHD. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

tabolism, abnormal intracellular Ca²⁺ handling, cardiocytes hypertrophy and apoptosis.

(1) Dysfunction of energy metabolism

During ischemia, cardiac energy metabolism is dramatically altered, resulting in the occurring of fatty acid oxidation as the dominant source of oxidative metabolism at the expense of glucose oxidation [68]. As shown in Fig. 6, the proteins in AMPK (5'-AMP-activated protein kinase) pathway such as carnitine palmitoyl transferase 1 (CPT-1), acetyl CoA carboxylase (ACC) and malonyl CoA decarboxylase (MCD) are important modifier factors in the energy metabolism of fatty acid [69].

Except with the imbalance of energy metabolism between fatty acid and glucose in ischemic heart, the other proteins related to energy such as F1F0 ATPase also become abnormal [70]. F1F0 ATPase is a critical enzyme to release ATP from catalytic F1 domain. But under ischemic conditions, this enzyme become an ATP hydrolase leading to an undesirable hydrolysis of ATP in ischemic heart [71, 72].

To overcome myocardial ischemia, endogenous reimbursement mechanism is activated. Hypoxia-inducible factor 1- α (HIF- α) may function as a master regulator of oxygen homeostasis [73, 74] By enhancing the transport of glucose, HIF- α increases the oxidation of pyruvate participating in the production of ATP in mitochondrion.

(2) Abnormal intracellular Ca²⁺ handling

Once heart ischemia happens, the concentration of norepinephrine/adrenaline, Ang II, ET-1, dopamine, acetylcholine and histamine in circulation significantly increases. By binding to their receptors on cardiocytes, these factors induce abnormal intracellular Ca²⁺ handling. In cardiocytes, both calcium release channel ryanodine receptor (RyR) and sarcoplasmic reticulum Ca²⁺-ATPase (SERCA2) can mediate Ca²⁺ releasing into the sarcoplasmic reticulum, thus inducing the decrease of intracellular Ca²⁺ concentration. L-type Ca²⁺ channels (LTCC) is a Ca²⁺ inflow channel which may induce the increase of intracellular Ca²⁺ concentration [75, 76] (Fig.6). The dysfunction of these proteins will directly result in abnormal Ca²⁺ handling.

(3) Hypertrophy of cardiocytes

Hypertrophy is another significant pathological change as the result of ischemia. Elevated Ca²⁺ concentration is an important factor for cardiocytes hypertrophy. High levels of Ang II and ET-1 after ischemia can seriously affect Ca²⁺ concentration in cardiocytes (Fig. 6). There are some other up-regulated cytokines contribute to cardiocytes hypertrophy. For example, both transforming growth factor beta-1 (TGF- β) and urotensin II can combine with their receptors on the surface of cardiocytes and then induce hypertrophy (Fig. 6).

(4) Cell apoptosis in ischemic heart

Except that the dysfunction of energy metabolism and abnormal intracellular Ca²⁺ handling can induce the apoptosis of cardiocytes, some other factors can also contribute to this process. As Fig. 6 shows, with the help of monoamine oxidase, 5-HT can produce H₂O₂. H₂O₂ and another noxious substances reactive oxygen species (ROS) both can lead to

DNA strand-breakage [77, 78]. Subsequently, poly(ADP-ribose) polymerase (PARP) induces inefficient cellular metabolic cycle and promote cardiocytes' death [79, 80].

On the other side, some endogenous factors such as thyroid hormone and adenosine are beneficial to cardiocytes in hypoxia condition. For example, adenosine is released by ischemic cardiocytes and subsequently reduces cellular injury and restores energetic homeostasis by binding to its receptors [81, 82]. The single pathway of adenosine and thyroid hormone are showed in Fig. 6.

3. Targets and corresponding medicines for CHD treatment

Proteins involved in the important CHD related pathways may be speculated as potential drug targets for CHD. Nowadays, more and more CHD targets have been identified and their corresponding medicines have received satisfactory effects for the treatment of CHD in clinic. These targets and their corresponding medicines are summarized in this section (Supplementary Table 2).

3.1 Anti hyperlipidemia

As the relationship between elevated plasma lipids and the development of atherosclerotic plaques has been well established, dyslipidemia has been considered as a major contributor to atherosclerosis-associated conditions such as CHD [83]. Nowadays, the most widely used medicine in clinic for anti hyperlipidemia is statins which is able to give a curative effect by inhibiting HMG-CoA reductase. As shown in Fig. 4, HMG-CoA reductase is an important enzyme for the formation of cholesterol. By inhibiting HMG-CoA reductase, the synthesis pathway of cholesterol can be blocked, thus decreasing the amount of LDL. Except statins, HMG-CoA synthase inhibitors, which is also in the pathway of cholesterol formation, is another important modifier for the down-regulation of LDL-c level [84] (Fig. 4). In addition, LDL-c can transform into ox-LDL in intima with the help of oxidases (Fig. 1). To decrease the effects induced by ox-LDL, some antioxidation medicines such as the inhibitors of xanthine oxidase (XO) [85, 86], NADPH oxidase [87] and myeloperoxidase (MPO) [88] have been discovered.

There are overwhelming evidences showing that a low plasma level of HDL is an important independent risk factor for CHD [83, 89]. Therapeutic intervention aimed at raising HDL has become a strategy increasingly adopted by Adult Treatment Panel III (ATP III) guidelines [90]. Niacin which stimulates ABCA1 to elevate the level of HDL has been widely used in clinic [91]. From pathways of HDL formation showed in Fig. 4, it can be seen that some other related proteins may be identified as potential drug targets. For example, several evidences have demonstrated that irritating the activity of LXR [30,92], PPAR [91,93] and scavenger receptor class B member 1 (SR-BI) [94] or inhibiting the activity of FXR and ARP1 [95, 96] can increase the level of HDL and

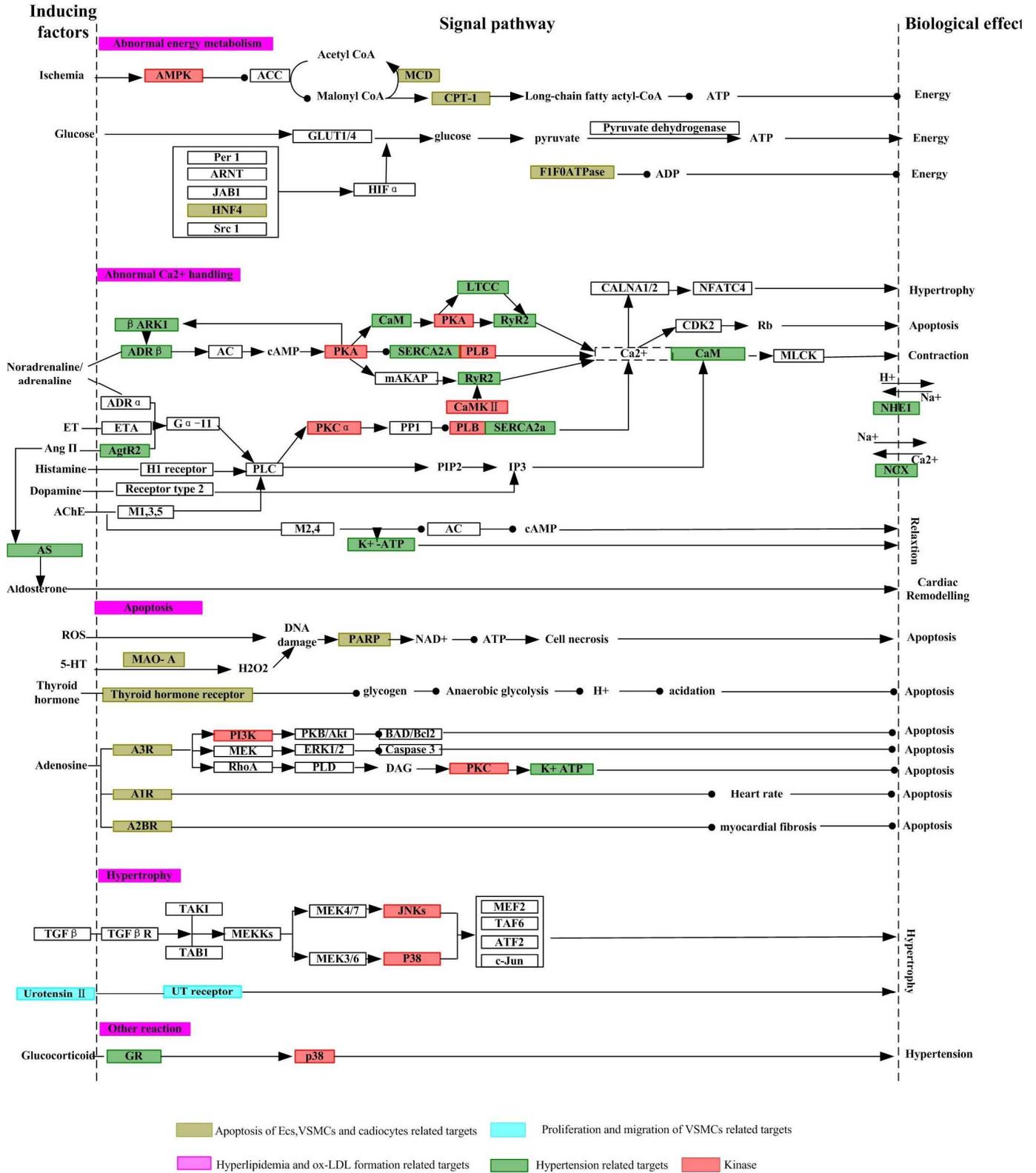


Fig. 6 The intracellular signal pathway of cardiocyte in CHD. All proteins were signed by rectangles. Drug targets were marked by different colors (the same as Table. 1) according to their different roles.

decrease the incidence of CHD obviously.

Decreasing the formation of cholesterol is another effective method for anti dyslipidemia. Bile acid sequestrants cholestyramine and colestipol promote the bile acid transformation from cholesterol, thus decreasing the LDL level and the morbidity of CHD [97]. The application of bile acid sequestrant in dyslipidemia has been recommended in ATP III. Cholesterol absorption inhibitor ezetimibe is another important anti hyperlipidemia medicine and widely used in clinic. By inhibiting the absorption of cholesterol, ezetimibe can efficiently control the level of LDL[83]. In addition, estrogen or hormone replacement therapy has been demonstrated their effects in favorable changes in lipid profiles, but its therapeutic effects are skeptical now because of the increased risk of thromboembolic events [98, 99].

3.2 Anti hypertension

Hypertension is another major risk factor involved in CHD. In ESH-ESC practice guidelines, five major classes of antihypertensive agents including ACE-inhibitors, angiotensin receptor blockers, β -adrenoreceptor inhibitors, calcium antagonists and thiazide diuretics are recommended for the initiation and maintenance of anti hypertensive treatment [100].

In the development of atherosclerosis and ischemic heart formation, RAS and sympathetic nervous system (SNS) are excited [101], thus increasing contractility of heart muscle and vascular smooth muscle and leading to high blood pressure. Inhibiting the activated RAS and SNS is an effective way to restrain the hemodynamic disturbance. Proteins playing important roles in the RAS pathway, such as ACE, AgtR and rennin, have been identified as drug targets (Fig. 3) and their corresponding medicines have received satisfactory effects in the treatment of hypertension [102-104]. Beta-adrenergic receptor kinase 1 (β AR kinase) and β -adrenoceptor (ADR β) in SNS pathway also have been identified as drug targets for inhibiting the contraction of cardiac muscle. ADR β antagonists also have been extensively adopted in clinic to inhibit the activity of SNS [105, 106].

As can be seen in Fig 6, inhibiting the elevated concentration of intracellular Ca²⁺ will be another effective method to decrease the contractility of heart muscle. Several important proteins involved in this biological process have been identified as drug targets, such as LTCC (L-type calcium channel), calmodulin (CaM), RyR2, SERCA2a, sodium/hydrogen exchanger 1 (NHE1) and NCX (sodium/calcium exchanger 1) [75,107-111]. The representative Ca²⁺ handling medicines have been shown in Supplementary Table 1.

Except regulating drug targets directly related to the pathways of hypertension, using diuretics to down-regulate high blood pressure indirectly is another effective method. Diuretics can inhibit the reabsorption of sodium at different segments of the renal tubular system. Through their effects on sodium and water balance, diuretics decrease ventricular stroke volume and cardiac output, finally leading to a fall in

arterial pressure. Diuretics have been used in the management of hypertension for approximately ninety years, and it is still widely applied in clinic now[112, 113].

In addition, ET-1 secreted by ECs after endothelial injury also induce the dysfunction of contractility [114] (Fig. 1 and Fig. 3). Endothelin receptor antagonists such as bosentan and tezosenta have been applied in clinic and received satisfied effects [115, 116]. Endothelin converting enzyme (ECE), a key enzyme for ET-1 formation (Fig. 2), is also considered as a potential drug target [117, 118].

3.3 Anti platelet treatment

Platelets play an important role in thrombus formation as well as in inflammatory reaction. American College of Chest Physicians (ACCP) evidence-based clinical practice guidelines (8th Edition) has supplied the intimate therapeutic methods for anti platelet in patients with CHD [119]. Alteplase, aspirin, clopidogrel, abciximab and vitamin K antagonist are main recommended medicines for CHD related condition. By regulating the different pathways in the process of platelet activation and aggravation (Fig. 5), these medicines receive therapeutic effects in clinic.

As shown in Fig. 5 and Supplementary Table 1, aspirin is the most representative anti platelet medicine through inhibiting cyclooxygenase. Clopidogrel and abciximab block the activation of platelet through inhibiting P2Y₁₂ receptor and GP IIb/IIIa, respectively. If patients have symptom of myocardial infarction, alteplase (plasminogen agonist) or vitamin K antagonist (thrombin/Factor II, Factor VII, Factor IX, and Factor X inhibitor) is available for thrombolysis therapy. In addition, some other medicines such as platelet-activating factor (PAF) acetylhydrolase agonists [120], proteinase activated receptor 1 (PAR-1) inhibitors [121] were also developed and some of them have been used in clinic.

3.4 Inhibiting inflammatory and immune reaction

As can be seen in Fig. 1, inflammatory and immune reaction coexisted in all stages of atherosclerosis. The development of inhibitors for inflammatory and immune reaction will be beneficial for the treatment of CHD. Some chemotactic factors such as MCP-1 and CCR2 have been identified as potential drug targets for inhibiting the inflammation in atherosclerosis. In immune reaction of atherosclerosis, complement system is activated. Complement C3/C5 convertase and complement C5 have been identified as drug targets and their corresponding medicines have been used in clinic[122].

Except the targets in extracellular, there are a lot of important intracellular targets contributing for the inflammation and immunity, such as NF-Kb [123] and RAGE [124, 125]. For example, as shown in Fig. 1, RAGEs bind the circulating AGEs, resulting in the generation of reactive oxygen species (ROS) and further activation of NF- κ B. To decrease the damage induced by AGEs, blocking the combination of AGE and RAGE becomes a new way to control CHD. RAGE

antagonist like ALT711 have showed their therapeutic effectiveness in clinic [126, 127]. Some other inflammation related drug targets and their representative medicines have been shown in Table. 1.

3.5 Inhibiting the apoptosis of ECs, VSMCs and cardiocytes

The apoptosis of ECs leads to the loss of integrity of endothelial monolayer and facilitates the migration and deposition of lipids or monocytes, thus propagating the atherosclerotic plaque development. The particles of apoptosis VSMCs not only are major components in atherosclerotic plaque but also contribute to destabilize the plaque. The apoptosis of cardiocytes, which is induced by heart ischemia, lead to the development of myocardial damage. Therefore, the treatment of apoptosis of ECs, VSMCs and cardiocytes are important ways to inhibit the process of CHD.

As inhibiting the activities of glutathione peroxidase and heme oxygenase can decrease the formation of ROS, thus avoiding injury of ECs, these proteins are considered as potential targets for the prevention of CHD in the future [128-130]. In addition, as eNOS agonists or NO donors such as S-nitrosothiols [131, 132], NONOates [133, 134] and tetrahydrobiopterin [135] can participate in NO formation, some of them have been used in clinic. As shown in Fig.2 and 3, TNF- α and some proteins in caspase pathway play a central role in the apoptosis of ECs and VSMCs, the inhibitors of caspases and TNF α have been under investigated and attained therapeutic effects through many years' clinic trails [136-138].

As Fig. 6 shows, proteins related to cardiocytes apoptosis have been summarized. Some of them have been identified as drug targets, such as adenosine receptor, PARP and monoamine oxidase A (MAO-A) [139-141]. In addition, Thyroid hormone can decrease the anaerobic glycolysis and inhibit the acidosis and apoptosis of cardiocytes. In clinic, the homolog of thyroid hormone such as DIPTA, GC-1 and its agonist desethylamiodarone have exhibited their therapeutic effects for the treatment of CHD [142-144].

Regulating proteins related to energy metabolism of cardiocytes is another way to restrain the apoptosis of cardiocytes. As Fig. 6 shows, CPT-1 and MCD, two important enzymes related to metabolism of fatty acid, have been identified as drug targets [145-147]. In addition, the activation of F1F0 ATP synthase can elevate content of ATP in myocardium, decrease energy consume and protect ischemic heart [70, 148, 149]. So F1F0 ATPase maybe a new potential drug target for controlling CHD.

Other drug targets and their representative medicines for inhibiting apoptosis of ECs, VSMCs and cardiocytes were listed in Table. 1.

3.6 Anti migration and proliferation of VSMCs

The proliferation and migration of VSMCs play an important role in the formation of atherosclerotic plaque. Fac-

tors like PDGF and MMPs are the most important cytokines which can directly induce proliferation and migration of VSMCs (Fig.1 and Fig. 3), while their antagonists such as trapidil (an antagonist of PDGF) [107] and doxycycline (an inhibitor of MMPs) [150] have been used as anti atherosclerotic medicines. Additionally, antagonists of serine/threonine-protein kinase mTOR (mTOR) [151, 152], arachidonate 5-lipoxygenase-activating protein (FLAP) [153] and urotensin-2 receptor (UT receptor) [154] also can inhibit migration and proliferation of VSMCs (Fig. 3). Other drug targets and their representative medicines for inhibiting migration and proliferation of VSMCs were listed in Table. 1.

3.7 Inhibiting the activity of CHD related kinases

Protein kinases play connective roles to transmit external stimuli to evoke intracellular biochemical reaction. Theoretically, they have comprehensive ability to regulate intracellular pathway and obtain desirable clinic effects. Kinases like protein kinase A (PKA), protein kinase C (PKC), phosphoinositide 3-kinase (PI3K), mitogen-activated protein kinase (JNK) and tyrosine protein kinase JAK (JAK) are the most common kinases related to CHD. Some kinases' inhibitors such as ruboxistaurin (a PKC- β inhibitor) [155] and wortmannin (a PI3K inhibitor) [156, 157] have been applied in clinic for the treatment of CHD. Though different therapeutic effects were received in clinic trials, the method searching for inhibitors of kinases to treat CHD still become a hot spot in drug discovery.

4. Medicine combination therapy for CHD

With the accumulation of experience in medicine application for the treatment of CHD, it has been found that monotherapy using only one compound can't make a satisfactory result because of complex pathogenesis of CHD. To overcome shortcomings of monotherapy, medicine combination therapy was adopted. Clinic studies have demonstrated that polytherapy using two or more medicines can obviously ameliorate therapeutic effects and elevate patients' quality of life.

4.1 Combination therapy for anti hyperlipidemia

Ezetimibe (a cholesterol absorption inhibitor) and statins (HMG-CoA reductase inhibitors) are both used to decrease LDL level. Coadministration offers an effective treatment option for patients with hypercholesterolemia, compared to statins monotherapy [158-160]. Colestipol, a bile acid sequestrant, is also used in combination therapy with statins. Trials have demonstrated that treatment with lovastatin and colestipol for 2-2.5 years slowed or reversed the progression of atherosclerosis, as assessed by angiography [161].

Combination therapy which can simultaneously decrease LDL level and elevate HDL level also used in clinic [162-164]. Combined therapy by statins and niacin (ABCA1 ago-

nist), which is recommended in ATP III guidelines, provides an option to help patients to attain their low-density lipoprotein cholesterol (LDL-C) goals and HDL goals [165, 166]. Monotherapy of probucol (SR-BI agonist) or fibrate (PPAR- α agonists) can raise HDL level. Actually, in clinic, the combination of probucol or fibrate with statins is used more often for the treatment of angina in CHD patients with hypercholesterolemia.

4.2. Combination therapy for anti hypertension

Diuretic, β -ADR antagonist, calcium antagonist, ACE inhibitor and Ang II receptor inhibitor are five mainly used medicines for anti-hypertension in clinic. According to 2007 ESH-ESC Practice Guidelines, any two categories can be simultaneously administrated [100].

ACE inhibitors are the first line medicines for the treatment of hypertension in the guideline of ESH/ESC. By coadministration with other anti-hypertension medicines, ACE inhibitors can receive a better therapeutic effect with lower toxicity [167-169]. The combined therapy with a diuretic and a β -ADR antagonist is also important in anti-hypertension. Their combination has been shown highly effective in reducing cardiovascular events in both diabetic and nondiabetic patients [170, 171]. The combination medicine including these two medicines has been proved by FDA as the first line anti-hypertension medicine. Except coadministration with an ACE inhibitor, Ca²⁺ channel antagonists can also simultaneously administrate with a β -ADR antagonist, an angiotensin receptor inhibitor or a diuretic. All the combined therapies have been demonstrated their therapeutic effects in clinic [172-174].

4.3. Combination therapy for anti platelet

Combination therapy for anti platelet in CHD treatment has been widely adopted in ACCP-8 guidelines [119]. Aspirin is a representative medicine for anti platelet therapy. It is usually coadmintrated with other anti platelet medicines such as clopidogrel (P2Y₁₂ receptor antagonist) to control the development of CHD in clinic [175, 176]. On the other side, ATP III guidelines also recommend the combination therapy of aspirin and statins for CHD patients to reduce prothrombotic state [90, 177, 178]. Five major clinical studies have demonstrated that the combination of pravastatin and aspirin was significantly more effective than each agent alone in reducing the relative risk of key cardiovascular endpoints including MI and ischemic stroke [158].

4.4. Other combination therapy in CHD treatment

High blood pressure combined with high cholesterol level constitutes a serious risk for CHD. Hence, decreasing LDL-C level and high blood pressure through coadministration of an antihypertensive and a statin has potential benefit in management of CHD [179, 180]. It was found that coadmin-

istration of valsartan and simvastatin was well tolerated and associated with significant reductions from baseline in blood pressure and LDL-C [181].

In the end stage of CHD, heart failure is a common event. To reverse this pathological processing, many combination therapies have been recommended. Published guidelines for the treatment of heart failure support the combination therapy by using an ACE inhibitor with a β -ADR inhibitor [182-184].

To simplify the effective treatment of CHD, polypill or fixed dose combinations (FDCs) have attracted the manufactures' favor [185]. In recent years, more and more FDCs such as Coveram (containing perindopril and amlodipine), Caduet (containing amlodipine and atorvastatin) and Vytorin (containing ezetimibe and simvastatin) have appeared on the market for the treatment of CHD [186-189]. Trials of a polypill containing a statin, three blood pressure lowering agents, aspirin and folic acid have been undertaken and shown exciting result in decreasing the incidence of CVDs by more than 80% after phase II clinical trial [190, 191].

5. Conclusion

CHD is a chronic and multifactorial disease. As hundreds of complex signal pathways containing in the CHD formation, it is important to construct biological network and protein database about CHD. In this review, we not only compiled CHD related signal pathways and protein database, but also provide a comprehensive knowledge of drug targets and their application in clinic for the treatment of CHD. As protein database include potential drug targets for the treatment of CHD, it may convenient the drug discovery in the future. With the development of drug discovery, it can be anticipated that more and more efficiently medicinal solutions would be applied in clinic for the treatment of CHD.

6. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/105/0>

Supplementary material includes: Supplementary Table 1 showing CHD related targets and their corresponding medicines ; Supplementary Table 2 listing 413 CHD proteins participating in multiple intracellular and extracellular signal pathways.

References

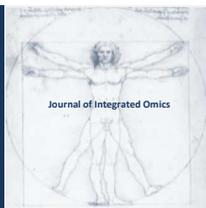
1. <http://www.who.int/mediacentre/factsheets/fs317/en/index.html>, Sep 10 2011.
2. <http://www.healthknowledge.org.uk/public-health-textbook/disease-causation-diagnostic/2b-epidemiology-diseases-phs/chronic-diseases/coronary-heart-disease>, Sep 27 2012.
3. F. Erling, Journal of the American College of Cardiology 47

- (2006) Suppl C:7-12. DOI: 10.1016/j.jacc.2005.09.068
4. B. Quang, P. Maxwell, L.W. Robert, *Int J Biochem Cell Biol* 41 (2009) 2109-2113. DOI: 10.1016/j.biocel.2009.06.002
 5. C.L. Andrew, K.G.Christopher, *Nature medicine* 8 (2002) 1235-1242. DOI: 10.1038/nm1102-1235
 6. L.R. Anna-Karin, K.H. Göran T, *Arterioscler Thromb Vasc Biol* 26 (2006) 2421-2432. DOI: 10.1161/01.ATV.0000245830.29764.84
 7. B. Ido, L.S. Francesca, *TRENDS in Immunology* 28 (2007) 360-365. DOI: 10.1016/j.it.2007.06.007
 8. N. Alexander, M.W. Cornelia, *Clin Immunol* 134 (2010) 25-32. doi:10.1016/j.clim.2009.05.006.
 9. Y.Q. Huo, F.L.Klaus, *TCM* 14 (2004) 18-22. DOI: 10.1016/j.tcm.2003.09.007
 10. H.T.W. Xander, R.M. Andrew, *Drug Discovery Today: Disease Mechanisms* 1 (2004) 9-15. d DOI: 10.1016/j.ddmec.2004.06.001
 11. R. Ross, J.A. Glomset, *Science* 180 (1973) 1332-1339. doi:10.1126/science.180.4093.1332
 12. U. Forstermann, T. Munzel, *Circulation* 113 (2006) 1708-1714. DOI: 10.1161/CIRCULATIONAHA.105.602532
 13. A. Tedgui, Z. Mallat, *Circ Res* 88 (2001) 877-887. DOI: 10.1161/hh0901.090440
 14. N. Mohamad, A.B. Judith, D.W. Andrew, *Arterioscler Thromb. Vasc. Biol.* 16 (1996) 831-842. DOI: 10.1161/01.ATV.16.7.831
 15. G.K. Hansson, *N Engl J Me.* 352 (2005) 1685-1695. DOI: 10.1056/NEJMra043430
 16. J.T. Willerson, D.J. Kereiakes, *Circulation* 108 (2003) 2060-2061. DOI: 10.1161/01.CIR.0000099580.72044.83
 17. M. Ruiz-Ortega, O. Lorenzo, M. Rupérez, V. Esteban, Y. Suzuki, S. Mezzano, J.J. Plaza, J. Egido, *Hypertension* 38 (2001) 1382-1387. DOI: 10.1161/hy1201.100589
 18. F. Brunner, C. BrasSilva, A.S. Cerdeira, A.F. LeiteMoreira, *Pharmacol Ther* 111 (2006) 508-531. DOI: 10.1016/j.pharmthera.2005.11.001
 19. J.L. Kingsley, W.L. Huff, K. Rust, A.M. Carroll, M.F. Martinez, G.E. Plopper, *Biochemical and Biophysical Research Communications* 293 (2002) 1000-1006. DOI: 10.1016/S0006-291X(02)00331-5
 20. L.L. Nguyen, P.A. Amore, *Int Rev Cytol* 204 (2001) 1-48. DOI: 10.1016/S0074-7696(01)04002-5.
 21. E.W. Raines, H. Koyama, N.O. Carragher, *Ann N Y Acad Sci* 902 (2000) 39-51. DOI: 10.1111/j.1749-6632.2000.tb06299.x
 22. R.E. Gilbert, S.A. Douglas, H. Krum, *Curr Opin Investig Drugs* 5 (2004) 276-282. PMID: 15083593
 23. Y.F. Qi, C.F. Xia, Y.H. Chen, L. Xue, Y.Z. Pang, C.S. Tang, *Chineses journal of pathophysiology* 18 (2001) 230-232. DOI: cnki:ISSN:1000-4718.0.2002-03-001
 24. M. Manuel, Q.B.Xu, *Experimental Gerontology* 36 (2001) 969-987. DOI: 10.1016/S0531-5565(01)00090-0
 25. M.M. Kockx, M.W. Knaapen, *J Pathol* 190 (2000) 267-280. DOI: 10.1002/(SICI)1096-9896(200002)190:3<267::AID-PATH523>3.0.CO;2-A
 26. R.B. Martin, J.B. Joseph, *Atherosclerosis* 138 (1998) 3-9. doi:10.1016/S0021-9150(98)00013-6
 27. J.R. Harrington, *Stem Cells* 18 (2000) 65-66. DOI: 10.1634/stemcells.18-1-65
 28. J.H. Qiao, J. Tripathi, N.K. Mishra, Y. Cai, S. Tripathi, X.P. Wang, S. Imes, M.C. Fishbein, S.K. Clinton, P. Libby, A.J. Lusis, T.B. Rajavashisth. *Am J Pathol* 150 (1997) 1687-1699. PMID: PMC1858194
 29. L. Esther, L. Dirk, B. Linda, D. Marjo, D. Mat, *TCM* 17 (2007) 118-123. DOI: 10.1016/j.tcm.2007.02.004
 30. N.B. Michelle, T. Peter, *Drug Discovery Today* 2 (2005) 97-103. DOI: 10.1016/j.ddstr.2005.05.018
 31. D.B. Grant, M.E. Ronald, *TRENDS in Endocrinology and Metabolism* 15 (2004) 158-15165. DOI: 10.1016/j.tem.2004.03.003
 32. L. Peter, *Nature* 420 (2002) 868-874. DOI: 10.1038/nature01323
 33. F. Mach, U. Schönbeck, P. Libby, *Atherosclerosis* 137 (1998) (Suppl):S89-S95. DOI: 10.1016/S0021-9150(97)00309-2
 34. U. Schöbeck, F. Mach, G.K. Sukhova, C. Murphy, J.Y. Bonnefoy, R.P. Fabunmi, P. Libby, *Circ Res* 81 (1997) 448-454. DOI: 10.1161/01.RES.81.3.448
 35. S. Stemme, B. Faber, J. Holm, O. Wiklund, J.L. Witztum, G.K.Hansson, *Proc Natl Acad Sci USA* 92 (1995) 3893-3897. PMID: PMC42068
 36. G. Wick, M. Romen, A. Amberger, B. Metzler, M. Mayr, G. Falkensammer, Q. Xu, *FASEB* 11 (1997) 1199-1207. PMID: 9367355
 37. S. Taleb, A.Tedgui, Z. Mallat, *J Intern Med* 263 (2008) 489-499. DOI: 10.1111/j.1365-2796.2008.01944.x
 38. Y.A. Mekori, D.D. Metcalfe, *J Allergy Clin Immunol* 104 (1999) 517-523. PMID: 10482820
 39. M. Adam, A.M. Yoseph, *IMAJ* 3 (2001) 216-221. PMID: 11987828
 40. L.K. Jim, S.C. David, A.A. Wael, S. Kelly, K. Guha, *Molecular Medicine Today* 6 (2000) 304-308. PMID: 10904247
 41. T.K. Petri, L. Miriam, J.L. Markus, A.L. Ken, *International Congress Series* 1262 (2004) 494-497. DOI: 10.1016/j.ics.2003.12.013
 42. L.K.L. Lee, P.T. Kovanen, *Arterioscler Thromb Vasc Biol* 12 (1992) 1329-1335. DOI: 10.1161/01.ATV.12.11.1329
 43. N. Di-Girolamo, D. Wakefield, *Dev Immunol* 7 (2000) 131-142. DOI: 10.1155/2000/82708
 44. N. Kanbe, A. Tanaka, M. Kanbe, A. Itakura, M. Kurosawa, H. Matsuda, *Eur J Immunol* 29 (1999) 2645-2649. DOI: 10.1002/(SICI)1521-4141(199908)29:08<2645::AID-IMMU2645>3.0.CO;2-1
 45. K. Maija, C.W. Allard, M.L. Chris, J.P. Jan, T.K. Karel, E.B. Anton, T.K. Petri, *J Am Coll Cardiol* 32 (1998) 606-612. DOI: 10.1016/S0735-1097(98)00283-6
 46. J.S. Sun, K.S. Galina, J.W. Paul, M. Yang, K. Shiro, L. Peter, A.M. Lindsey, M.C. Jon, G.P. Shi, *Nat Med* 13 (2007) 719-724. DOI: 10.1038/nm1601
 47. M.J. Leskinen, H.M. Heikkilä, M.Y. Speer, J.K. Hakala, M. Laine, P.T. Kovanen, K.A. Lindstedt, *Exp Cell Res* 312 (2006) 1289-1298. DOI: 10.1016/j.yexcr.2005.12.033
 48. R.S. Lord, Y.V. Bobryshev, *Atherosclerosis* 146 (1999) 197-198. DOI: 10.1016/S0021-9150(99)00119-7
 49. Y.V. Bobryshev, R.S. Lord, S.P. Rainer, V.F. Munro, *Acta Histochem* 98 (1996) 185-194. DOI: 10.1016/S0065-1281(96)80037-7
 50. D. Alvarez, E.H. Vollmann, U.H. von Andrian, *Immunity* 29 (2008) 325-342. DOI: 10.1016/j.immuni.2008.08.006
 51. K. Shortman, S.H. Naik, *Nat Rev Immunol* 7 (2007) 19-30. DOI: 10.1038/nri1996
 52. C.J. Alderman, P.R. Bunyard, B.M. Chain, J.C. Foreman, D.S. Leake, D.R. Katz, *Cardiovasc Res* 55 (2002) 806-819. DOI: 10.1016/S0008-6363(02)00447-9
 53. X. Zhang, A. Niessner, T. Nakajima, W. Ma-Krupa, S.L. Kopecy, R.L. Frye, J.J. Goronzy, C.M. Weyand, *Circ Res* 98 (2006) 524-531. DOI: 10.1161/01.RES.0000204452.46568.57
 54. M.R. Zaverio, *Nature Medicine* 8 (2002) 1227-1234. DOI: 10.1038/nm1102-1227
 55. G. Meinrad, *Blood Cells, Molecules, and Diseases* 36 (2006) 206-210. DOI: 10.1016/j.bcmd.2005.12.022

56. S. Massberg, M. Gawaz, S.Gruner, V. Schulte, I. Konrad, D. Zohlhofer, U. Heinzmann, B. Nieswandt, *J Exp Med* 197 (2003) 41-49. DOI: 10.1084/jem.20020945
57. M. Arya, J.A. Lopez, G.M. Romo, M.A. Cruz, A. Kasirer-Friede, S.J. Shattil, B. Anvari, *J Thromb Haemost* 1 (2003) 1150-1157. PMID: 12871313
58. M.L. Kahn, *Semin Thromb Hemost* 30 (2004) 419-425. DOI: 10.1055/s-2004-833477
59. S.R. Coughlin, *Nature* 407 (2000) 258-264. DOI: 10.1038/35025229
60. L. Covic, A.L. Gresser, A. Kuliopulos, *Biochemistry* 39 (2000) 5458-5467. DOI: 10.1021/bi9927078
61. C. Gachet, *Ann Med* 32 (2000) 15-20. PMID: 11209975
62. B. Savage, M. Cattaneo, Z.M. Ruggeri, *Curr Opin Hematol* 8 (2001) 270-276. PMID: 11604561
63. V. Henn, J.R. Slupsky, M. Gräfe, I. Anagnostopoulos, R. Förster, G. Müller-Berghaus, R.A. Kroczeck, *Nature* 91 (1998) 591-594. DOI: 10.1038/35393
64. M. Gawaz, K. Brand, T. Dickfeld, G. Pogatsa-Murray, S. Page, C. Bogner, W. Koch, A. Schomig, F. Neumann, *Atherosclerosis* 148 (2000) 75-85. PMID: 10580173
65. A. Schober, D. Manka, P. Hundelshausen, Y. Huo, P. Hanrath, I.J. Sarembock, K. Ley, C. Weber, *Circulation* 106 (2002) 1523-1529. DOI: 10.1161/01.CIR.0000028590.02477.6F
66. P. Hundelshausen, R.R. Koenen, M. Sack, S.F. Mause, W. Adriaens, A.E. Proudfoot, T.M. Hackeng, C. Weber, *Blood* 105 (2005) 924-930. DOI: 10.1182/blood-2004-06-2475
67. R. Ross, D.F. Bowen-Pope, E.W. Raines, *Ann N Y Acad Sci* 454 (1985) 254-260. DOI: 10.1111/j.1749-6632.1985.tb11865.x
68. J.T. Whitmer, J.A. Idell-Wenger, M.J. Rovetto, J.R. Neely, *J Biol Chem* 253 (1978) 4305-4309. PMID: 659417
69. R.U. John, D.L. Gary, *Basic Res Cardiol* 104 (2009) 203-210. DOI: 10.1007/s00395-009-0003-9
70. J.G. Gary, A.M. Palma, K. Lee, S.Y. Donna, *The International Journal of Biochemistry and Cell Biology* 40 (2008) 2698-2701. DOI:10.1016/j.biocel.2008.06.013
71. R.H. Fillingame, W. Jiang, O.Y. Dmitriev, *Journal of Experimental Biology* 203 (2000) 9-17. PMID: 10600668
72. M.R. Duchon, *Molecular Aspects of Medicine* 25 (2004) 365-451. DOI: 10.1016/j.mam.2004.03.001
73. K.W.T. Kenneth, L.E. Huang, *J Biol Chem* 280 (2005) 38102-38107. DOI: 10.1074/jbc.M504342200
74. Z. Oliver, F.S. Thomas, E. Thomas, W. Alexander, F.F. Martin, *Biochemical and Biophysical Research Communications* 376 (2008) 315-320. DOI: 10.1016/j.bbrc.2008.08.152
75. F.D. Angela, A.B. Nicole, P. Pierre, G.C. Marco, *Pharmacology and Therapeutics* 113 (2007) 247-263. DOI: 10.1016/j.pharmthera.2006.08.007
76. J.B. Ivor, D.S. Michael, *J Clin Invest* 115 (2005) 495-499. DOI: 10.1172/JCI200524477
77. Y. Shimizu, S. Minatoguchi, K. Hashimoto, Y. Uno, M. Arai, N. Wang, X. Chen, C. Lu, G. Takemura, M. Shimomura, T. Fujiwara, H. Fujiwara, *J Am Coll Cardiol* 40 (2002) 1347-1355. DOI: 10.1016/S0735-1097(02)02158-7
78. A. Maurel, C. Hernandez, O. Kunduzova, G. Bompert, C. Cambon, A.Parini, B.Frances, *Am J Physiol Heart Circ Physiol* 284 (2003) 1460-1467. DOI: 10.1152/ajpheart.00700.2002
79. M. Chen, Z. Zsengeller, C.Y. Xiao, C. Szabó, *Cardiovasc Res* 63 (2004) 682-688. DOI: 10.1016/j.cardiores.2004.04.018
80. C. Szabó, *Pharmacological Research* 52 (2005) 34-43. DOI: 10.1016/j.phrs.2005.02.017
81. H.T. Sommerschild, K.A. Kirkeboen, *Acta Anaesthesiol Scand* 44 (2000) 1038-1055. DOI: 10.1034/j.1399-6576.2000.440903.x
82. J.P. Headrick, B. Hack, K.J. Ashton, *Am J Physiol Heart Circ Physiol* 285 (2003) 1797-1819. DOI: 10.1152/ajpheart.00407.2003
83. S.J. Kishor, M.K. Kathiravan, S.S. Rahul, *Bioorganic and Medicinal Chemistry* 15 (2007) 4674-4699. DOI: 10.1016/j.bmc.2007.04.031
84. A. Carazo, M.J. Alejandre, A. Louktibi, A. T. Linares, *Biochem J* 334 (1998) 113-119. PMID: PMC1219669
85. U.E. Ekelund, R.W. Harrison, O. Shokek, R.N. Thakkar, R.S. Tunin, H. Senzaki, D.A. Kass, E. Marbán, J.M. Hare, *Circ Res* 85 (1999) 437-445. DOI: 10.1161/01.RES.85.5.437
86. K. M. Minhas, R.M. Saraiva, K.H. Schuleri, S. Lehrke, M. Zheng, A.P. Saliaris, C.E. Berry, L.A. Barouch, K.M. Vandegaer, D. Li, J.M. Hare, *Circ Res* 98 (2006) 271-279. DOI: 10.1161/01.RES.0000200181.59551.71
87. J.A. Ellis, S. Mayer, O.T. Jones, *Biochem J* 251 (1988) 888-891. PMID: PMC1149085
88. Heart Protection Study Collaborative Group, *Lancet* 360 (2002) 7-22. DOI: 10.1016/S0140-6736(02)09327-3
89. M.J. Chapman, G. Assmann, J.C. Fruchart, J. Shepherd, C. Sirtori, *Curr Med Res Opin* 20 (2004) 1253-1268. PMID: 15324528
90. M.G. Scott, NHLBI produced publications, Third report of the expert panel on detection, evaluation, and treatment of high blood cholesterol in adults, 2002.
91. E. Arnold von, *Drug Discovery Today* 1 (2004) 177-187. DOI: 10.1016/j.ddstr.2004.09.004
92. T. Peter, J.M. David, *Molecular Endocrinology* 17 (2003) 985-999. doi: 10.1210/me.2003-0061
93. Z. Fokko, P. Jorge, *Biochimica et Biophysica Acta* 1771 (2007) 972-982. DOI: 10.1016/j.bbaliip.2007.04.021
94. M. Ishigami, S. Yamashita, N. Sakai, K.I. Hirano, T. Arai, T. Maruyama, S. Takami, M. Koyama, K. Kameda-Takemura, Y. Matsuzawa, *European Journal of Clinical Investigation* 27 (1997) 285-292. DOI: 10.1046/j.1365-2362.1997.1040657.x
95. A.D. Mooradian, M.J. Hass, N.C. Wong, *Diabetes* 53 (2004) 513-520. DOI: 10.2337/diabetes.53.3.513
96. C. Thierry, S. Ekkehard, D. Hélène, P.T. Inés, S. Audrey, K. Vladimir, F. Jean-Charles, D. Jean, H.W. Dean, K. Folkert, S. J. Bart, *Clin Invest* 109 (2002) 961-971. DOI: 10.1172/JCI0214505
97. E.A. Stein, K.W. Heimann, *Sa Mediese Tydskrif* 1975 1252-1256. PMID: 168650
98. M.H. Davidson, K.C. Maki, S.K. Karp, K.A. Ingram, *Drugs and Aging* 19 (2002) 169-178. PMID: 12027776
99. L.H. Kuller, *Arterioscler Thromb Vasc Biol* 23 (2003) 11-16. DOI: 10.1161/01.ATV.0000046033.32478.6D
100. M. Giuseppe, D.B. Guy, D. Anna, C. Enata, F. Robert, G. Giuseppe, G. Guido, M.H. Anthony, E.K. Sverre, L. Stephane, *Journal of Hypertension* 25 (2007) 1751-1762. DOI: 10.1093/eurheartj/ehm236
101. M.K. David, K. Henry, *Nature drug discovery* 6 (2007) 127-139. DOI: 10.1038/nrd2219
102. E.K. Mohammed, W.B. Abul, J. Edward, A.A. Imad, *Journal of the American College of Cardiology* 37 (2001) 1757-1764. DOI: 10.1016/S0735-1097(01)01229-3
103. M. Packer, R.M. Califf, M.A. Konstam, H. Krum, J.J. McMurray, J.L. Rouleau, K. Swedberg, *Circulation* 106 (2002) 920-926. DOI: 10.1161/01.CIR.0000029801.86489.50
104. J.M. Wood, J. Maibaum, J. Rahuel, M.G. Grütter, N.C. Cohen, V.Rasetti, H.Rüger, R.Göschke, S.Stutz, W.Fuhrer, *Biochem Biophys Res Commun* 308 (2003) 698-705. DOI: 10.1016/S0006-291X(03)01451-7
105. T. Noboru, *Pharmacology and Therapeutics* 100 (2003) 215-234. DOI: 10.1016/j.pharmthera.2003.09.001

106. F. Waagstein, M.R. Bristow, K. Swedberg, F. Camerini, M.B. Fowler, M.A. Silver, E.M. Gilbert, M.R. Johnson, F.G. Goss, A. Hjalmarson, *Lancet* 342 (1993) 1441-1446. DOI: 10.1016/0140-6736(93)92930-R
107. J.D. Victor, C.B. Ruediger, G.S. Daniel, *Nature medicine* 8 (2002) 1249-1256. DOI: 10.1038/nm1102-1249
108. A.A. James, E.W. Christine, X. Qi, *Lancet* 356 (2000) 1287-1289. DOI: 10.1016/S0140-6736(00)04440-8
109. P. Kirchhof, L. Fabritz, A. Kilic, F. Begrow, G. Breithardt, M. Kuhn, *J Mol Cell Cardiol* 36 (2004) 691-700. DOI: 10.1016/j.jmcc.2004.03.007
110. M. Yano, S. Kobayashi, M. Kohno, M. Doi, T. Tokuhisa, S. Okuda, M. Suetsugu, T. Hisaoka, M. Obayashi, T. Ohkusa, *Circulation* 107 (2003) 477-484. DOI: 10.1161/01.CIR.0000044917.74408.BE
111. F. del Monte, S.E. Harding, G.W. Dec, J.K. Gwathmey, R.J. Hajjar, *Circulation* 105 (2002) 904-907. DOI: 10.1161/hc0802.105564
112. O. Suzanne, *Hypertension* 41 (2003) 1006-1009. DOI: 10.1161/01.HYP.0000070905.09395.F6
113. J.D. Curb, S.L. Pressel, J.A. Cutler, P.J. Savage, W.B. Applegate, H. Black, G. Camel, B.R. Davis, P.H. Frost, N. Gonzalez, *JAMA* 276 (1996) 1886-1892. DOI: 10.1001/jama.1996.03540230036032
114. M.P. Love, J.J. McMurray, *Drugs Aging* 18 (2001) 1425-1440. PMID: 11419917
115. L.E. Spieker, G. Noll, F.T. Ruschitzka, T.F. Luscher, *J Am Coll Cardiol* 37 (2001) 1493-1505. DOI:10.1016/S0735-1097(01)01210-4
116. S.H. Ellahham, V. Charlon, Z. Abassi, K.A. Calis, W.K. Choucair, *Clin Cardiol* 23 (2000) 803-807. DOI: 10.1002/clc.4960231128
117. P.J. Cowburn, J.G.F. Cleland, *Eur Heart J* 22 (2001) 1772-1784. DOI: 10.1053/euhj.2000.2557
118. M.P. Love, W.G. Haynes, G.A. Gray, D.J. Webb, J.J.V. McMurray, *Circulation* 94 (1996) 2131-2137. DOI: 10.1161/01.CIR.94.9.2131
119. H. Jack, G. Gordon, W.A. Gregory, H. Robert, J.S. Holger, *Chest* 133 (2008) 71-109. DOI:10.1378/chest.08-0693
120. H.M. Garcia, P.W. Serruys, *Curr Opin Lipidol* 20 (2009) 327-332 DOI: 10.1097/MOL.0b013e32832dd4c7.
121. R.C. Becker, D.J. Moliterno, L.K. Jennings, K.S. Pieper, J. Pei, A. Niederman, K.M. Ziada, G. Berman, J. Strony, D. Joseph, *Lancet* 373 (2009) 919-928. DOI: 10.1016/S0140-6736(09)60230-0
122. C.B. Granger, K.W. Mahaffey, W.D. Weaver, *Circulation* 108 (2003) 1184-1190. DOI: 10.1161/01.CIR.0000087447.12918.85
123. O. Kutuk, H. Basaga, *Trends Mol Med* 9 (2003) 549-557. DOI: 10.1016/j.molmed.2003.10.007
124. G. Basta, *Atherosclerosis* 196 (2008) 9-21. DOI: 10.1016/j.atherosclerosis.2007.07.025
125. A.P. Burke, F.D. Kolodgie, A. Zieske, D.R. Fowler, D.K. Weber, P.J. Varghese, *Arterioscler Thromb Vasc Biol* 24 (2004) 1266-1271. DOI: 10.1161/01.ATV.0000131783.74034.97
126. C. Riccardo, M.F. Josephine, C.T. Merlin, T. Vicki, G.D. Rachael, C.B. Wendy, *Circ Res* 92 (2003) 785-792. DOI: 10.1161/01.RES.0000065620.39919.20
127. D.A. Kass, E.P. Shapiro, M. Kawaguchi, *Circulation* 104 (2001) 1464-1470. DOI: 10.1161/hc3801.097806
128. F.L. Muller, M.S. Lustgarten, Y. Jang, A. Richardson, H. Van Remmen, *Free Radic Biol Med* 43 (2007) 477-503. DOI: 10.1016/j.freeradbiomed.2007.03.034
129. B. Pascale, K. Oxana, M. Emanuela, C. Claudie, B. Daniele, R. Laura, H.S. Marie, N. Silvia, L. Nathalie, P. W. C. Angelo, *Circulation* 112 (2005) 3297-3305. DOI: 10.1161/CIRCULATIONAHA.104.528133
130. S. Brian, J.E. Zuckerbraun, H. Andrew, S. Said, T. Edith, *Journal of Surgical Research* 124 (2005) 256-263. DOI: 10.1016/j.jss.2004.10.022
131. M.A. Moro, V.M. Darley-Usmar, I. Lizasoain, Y. Su, R.G. Knowles, M.W. Radomski, S. Moncada, *Br J Pharmacol* 116 (1995) 1999-2004. PMID: PMC1908945
132. J. Loscalzo, *Circ Res* 88 (2001) 756-762. DOI: 10.1161/hh0801.089861
133. W.L. David, *Expert Opinion on Therapeutic Patents* 11 (2001) 999-1005. PMID: 22998044
134. Y. Tadachiko, J.B. Richard, *Proceedings of the Society for Experimental Biology and Medicine* 225 (2000) 200-206. PMID: 11082214
135. W.C.F. Maier, R.B. Lutolf, M. Fleisch, C. Seiler, O.M. Hess, B. Meier, T.F. Luscher, *J Cardiovasc Pharmacol* 35 (2000) 173-178. PMID: 10672847
136. J.C. Randle, M.W. Harding, G. Ku, *Exp Opin Invest Drugs* 10 (2001) 1207-1209. DOI: 10.1517/13543784.10.7.1207
137. B. Bozkurt, G. Torre-Amione, M.S. Warren, J. Whitmore, O.Z. Soran, A.M. Feldman, D.L. Mann, *Circulation* 103 (2001) 1044-1047. DOI: 10.1161/01.CIR.103.8.1044
138. L.M. Douglas, J.V.M. John, P. Milton, S. Karl, S.B. Jeffrey, S.C. Wilson, *Circulation* 109 (2004) 1594-1602. DOI: 10.1161/01.CIR.0000124490.27666.B2
139. Z.G. Gao, K.A. Jacobson, *Expert Opinion on Emerging Drugs* 12 (2007) 479-492. DOI: 10.1517/14728214.12.3.479
140. C. Szabó, *Pharmacol Res* 52 (2005) 34-43. DOI: 10.1016/j.phrs.2005.02.017
141. P. Bianchi, O. Kunduzova, E. Masini, C. Cambon, D. Bani, L. Raimondi, M.H. Seguelas, S. Nistri, W. Colucci, N. Leducq, A. Parini, *Circulation* 112 (2005) 3297-305. DOI: 10.1161/CIRCULATIONAHA.104.528133
142. E. Morkin, G.D. Pennock, P.H. Spooner, J.J. Bahl, S. Goldman, *Thyroid* 12 (2002) 527-533. DOI: 10.1152/ajpheart.01293.2008
143. S.U. Trost, E. Swanson, B. Gloss, D.B. Wang-Iverson, H. Zhang, *Endocrinology* 141 (2000) 3057-3064. DOI: 10.1210/en.141.9.3057
144. O. Bakker, H.C. van Beeren, W.M. Wiersinga, *Endocrinology* 134 (1994) 1665-1670. DOI: 10.1210/en.134.4.1665
145. M.P. Chandler, P.N. Chavez, T.A. McElfresh, H. Huang, C.S. Harmon, W.C. Stanley, *Cardiovasc Res* 59 (2003) 143-151. DOI: 10.1016/S0008-6363(03)00327-4
146. J.F. Cheng, M. Chen, D. Wallace, S. Tith, M. Haramura, B. Liu, *J Med Chem* 49 (2006) 1517-1525. DOI: 10.1021/jm050109n
147. J.F. Cheng, Y. Huang, R. Penuliar, M. Nishimoto, L. Liu, T. Arrhenius, *J Med Chem* 49 (2006) 4055-4058. DOI: 10.1021/jm0605029
148. K.S. Atwal, P. Wang, W.L. Rogers, P. Slep, H. Monshizadegan, F.N. Ferrara, *Journal of Medicinal Chemistry* 47 (2004) 1081-1084. DOI: 10.1021/jm030291x
149. G.J. Grover, K.S. Atwal, P.G. Slep, F.L. Wang, H. Monshizadegan, T. Monticello, *American Journal of Physiology* 287 (2004) 1747-1755. DOI: 10.1152/ajpheart.01019.2003
150. H. Chen, D. Li, G.J. Roberts, T. Saldeen, J.L. Mehta, *Cardiovasc Res* 59 (2003) 7-13. DOI: 10.1016/S0008-6363(03)00349-3
151. W. Ron, P. Rajbabu, S.B. Mary, P.G. Cindy, L. Laurent, F. Jana, W. Roswitha, H. David, *Cardiovascular Radiation Medicine* 4 (2003) 34-38. DOI: 10.1016/S1522-1865(03)00121-5
152. S.R. Tetsuo, M. Julie, T. Oleg, C. Kimber, C.S. Megan, J.M. Warren, I. Seigo, *Circulation* 107 (2003) 1664-1670. DOI:

- 10.1161/01.CIR.0000057979.36322.88
153. H. Hakonarson, S. Thorvaldsson, A. Helgadóttir, D. Gudbjartsson, F. Zink, M. Andresdóttir, A. Manolescu, D.O. Arnar, K. Andersen, A. Sigurdsson, G. Thorgeirsson, A. Jonsson, U. Agnarsson, H. Bjornsdóttir, G. Gottskalksson, A. Einarsson, H. Gudmundsdóttir, A.E. Adalsteinsdóttir, K. Gudmundsson, K. Kristjánsson, T. Hardarson, A. Kristinsson, E.J. Topol, J. Gulcher, A. Kong, M. Gurney, G. Thorgeirsson, K. Stefansson, *JAMA* 293 (2005) 2245-2256. DOI: 10.1001/jama.293.18.2245.
 154. M. Tölle, M.vander Giet, *Peptides* 29 (2008) 743-763. DOI: 10.1016/j.peptides.2007.08.029
 155. A.J. Boyle, D.J. Kelly, Y. Zhang, A.J. Cox, R.M. Gow, K. Way, S. Itescu, H. Krum, R.E. Gilbert, *J Mol Cell Cardiol* 39 (2005) 213-221. DOI: 10.1016/j.yjmcc.2005.03.008
 156. C.J. Vlahos, S.A. McDowell, A. Clerk, *Nat Rev Drug Discov* 2 (2003) 99-113. DOI: 10.1038/nrd1009
 157. T. Shioi, P.M. Kang, P.S. Douglas, J. Hampe, C.M. Yballe, J. Lawitts, *EMBO J* 19 (2000) 2537-2548. DOI: 10.1093/emboj/19.11.2537
 158. J.M. CruzF, G.V. Bedarida, J. Adgey, C. Allen, A.O. JohnsonL, R. Massaad, *Int J Clin Pract* 59 (2005) 619-627. DOI: 10.1111/j.1368-5031.2005.00565.x
 159. M. Farnier, M. Volpe, R. Massaad, M.J. Davies, C. Allen, *Int J Cardiol* 102 (2005) 327-332. DOI: 10.1016/j.ijcard.2005.01.022
 160. J. McKenney, C.M. Ballantyne, T.A. Feldman, W.E. Brady, A. Shah, M.J. Davies, J. Palmisano, Y.B. Mitchel, *Med Gen Med* 7 (2005) 3. PMID: PMC1681670
 161. A. Frisinghelli, A. Mafriaci, *Clin Drug Investig* 27 (2007) 591-604. PMID:17705568
 162. M.H. Davidson, *Expert Opin Drug Saf* 5 (2006) 145-156. DOI: 10.1517/14740338.5.1.145
 163. M. Farnier, M.W. Freeman, G. Macdonnell, I. Perevozskaya, M.J. Davies, Y.B. Mitchel, B. Gumbiner, *Eur Heart J* 26 (2005) 897-905. DOI: 10.1093/eurheartj/ehi231
 164. V. Barrios, N. Amabile, F. Paganelli, J.W. Chen, C. Allen, A.O. Johnson, R. Massaad, K. Vandormael, *Int J Clin Pract* 59 (2005) 1377-1386. DOI: 10.1111/j.1368-5031.2005.00714.x
 165. D.R. Levy, T.A. Pearson, *Clin Cardiol* 28 (2005) 317-320. PMID:16075823
 166. J.M. McKenney, P.H. Jones, H.E. Bays, R.H. Knopp, M.L. Kashyap, G.E. Ruoff, M.E. McGovern, *Atherosclerosis* 192 (2007) 432-437. DOI: 10.1016/j.atherosclerosis.2006.11.037
 167. K. Jamerson, M.A. Weber, G.L. Bakris, D. Björn, P. Bertram, S. Victor, *N Engl J Med* 359 (2008) 2417-2428. DOI: 10.1056/NEJMoa0806182
 168. J.M. Neutel, D.H. Smith, M.A. Weber, L. Schofield, D. Purkayastha, M. Gatlin, *Clin Hypertens* 7 (2005) 641-646. DOI: 10.1111/j.1524-6175.2005.04615
 169. M.L. Otero, *Vascular health and risk management* 3 (2007) 255-263. PMID: PMC2293964
 170. R.G. Sehlienger, M.E. Kraenzlin, S.S. Jiek, R.M. Christoph, *JAMA* 292 (2004)1326-1332. DOI: 10.1001/jama.292.11.1326
 171. M. Moser, J. Setaro, *Med Clin North Am* 88 (2004) 167-187. PMID: 14871058
 172. H.T. Ong, *BMJ* 334 (2007) 946-949. DOI: 10.1136/bmj.39185.440382.47
 173. E. Pimenta, S. Oprel, *Vasc Health Risk Manag* 4 (2008) 653-664. PMID: PMC2515425
 174. X.B. Xie, J. Hong, S. Tao, N.F. Li, *J Clin Cardiol* 25 (2009) 627-628. PMID: PMC338329
 175. E.L. Eisenstein, K.J. Aanstrom, D.F. Kong, L.K. Shaw, R.H. Tuttle, D.B. Mark, *JAMA* 297 (2007) 159-168. DOI: 10.1001/jama.2011.551
 176. N. Horst, G. Bülent, H. Christoph, S. Martin, M. Andreas, *Eur Heart J* 24 (2003) 1744-1749. DOI: 10.1016/S0195-668X(03)00442-1
 177. M. Pignone, S. Earnshaw, J.A.Tice, M.J. Pletcher, *Ann Intern Med* 144 (2006) 326-336. PMID: 16520473
 178. D.S. Wald, G. Morton, K. Walker, N. Iosson, N.P. Curzen, *Ann Pharmacother* M 41 (2007) 1644-1647. DOI: 10.1345/aph.1K232
 179. K. Ratheiser, J. Dusleag, K. Seilt, G. Titsche, W. Klein, *Clin Cardiol* 15 (1992) 647-654. PMID: 1395199
 180. V.G. Athyros, D.P. Mikhailidis, A.A. Papageorgiou, V.I. Bouloukos, A.N. Pehlivanidis, A.N. Symeonidis, M. Elisaf, *J Hum Hypertens* 18 (2004) 781-788. DOI: 10.1038/sj.jhh.1001748
 181. L.C. Rump, E. Baranova, B. Okopien, M. Weisskopf, A. Kandra, P. Ferber, *Clin Ther* 30 (2008) 1782-1793. DOI: 10.1016/j.clinthera.2008.10.004
 182. C. John, D. Henry, D. Helmut, F. Ferenc, K. Michel, *Eur Heart J* 26 (2005) 1115-1140. DOI: 10.1093/eurheartj/ehi204
 183. W.P. Abhayaratna, W.T. Smith, N.G. Becker, T.H. Marwick, I.M. Jeffery, D.A. McGill, *Med J Aust* 184 (2006) 151-154. PMID:16489896
 184. L. Daniel, K. Satish, G.L. Martin, J.B. Emelia, J.K. Michelle, K.L. Kalon, M.M. Joanne, S.V. Ramachandran, *N Engl J Med* 347 (2002) 1397-1402. PMID:12409541
 185. K. Vijay, P. Bhagwat, *Drug Discovery Today* 5 (2008) 63-71. DOI: 10.1038/ncpcardio1424
 186. A.J. Scheen, J.M. Krzesinski, *Rev Med Liege* 64 (2009) 223-227. PMID: 19514543
 187. K. Mckeage, M. Asif, A. Siddiqui, *American journal of cardiovascular drugs* 8 (2008) 51-67. DOI: 10.2165/0129784-200808010-00007
 188. P.C. Christopher, P.G. Robert, A.B. Michael, A.H. Robert, L.P. John, M.S. Christine, S. John, A.M. Thomas, H.M. Carolyn, V. Enrico, B. Eugene, M.C. Robert, *American Heart Journal* 156 (2008) 826-832. PMID: 19061694
 189. M.B. Christie, A. Nicola, Y. Zhong, R.K. Thomas, P. Joanne, *American Heart Journal*. 149 (2005) 464-473. PMID: 15864235
 190. S. Yusuf, P. Pais, R. Afzal, D. Xavier, K. Teo, J. Eikelboom, A. Sigamani, V. Mohan, R. Gupta, N. Thomas, *Lancet* 373 (2009) 1341-1351. DOI: 10.1016/S0140-6736(09)60611-5
 191. N.J. Wald, M.R. Law, *BMJ* 326 (2003) 1419-1428. DOI: 1007-9742.0.2004-01-016



REVIEW ARTICLE | DOI: 10.5584/jiomics.v2i2.111

Organelle proteomics in skeletal muscle biology

Kay Ohlendieck*

Muscle Biology Laboratory, Department of Biology, National University of Ireland, Maynooth, Co. Kildare, Ireland

Received: 18 September 2012 Accepted: 17 October 2012 Available Online: 04 November 2012

ABSTRACT

The cell biological profile of skeletal muscle tissues is highly complex and variable due to the molecular heterogeneity and cellular plasticity of contractile fibres and their supportive structures. Mass spectrometry-based proteomics has been used to study global changes in muscle during maturation, differentiation and physiological adaptations, as well as following pathological insults. However, due to the dynamic protein expression range of contractile cells, the findings from large-scale biochemical surveys of crude tissue extracts were limited to mostly soluble and abundant protein species. To overcome this technical problem, organelle proteomics was applied to study distinct subcellular fractions from skeletal muscle preparations. Tissue pre-fractionation procedures significantly reduce sample complexity and thus allow a more comprehensive cataloging of highly complex protein mixtures. This article reviews the impact of recent subproteomic studies of skeletal muscle and discusses findings from changes in the proteome of mitochondria, surface membranes, sarcoplasmic reticulum, cytosol and the contractile apparatus in normal, transforming and pathological muscle.

Keywords: Contractile apparatus; Mitochondria; Muscle proteomics; Sarcolemma; Sarcoplasmic reticulum.

1. Introduction

Comparative proteomics presents one of the most powerful analytical tools to determine global changes in distinct protein constellations, including the establishment of alterations in protein abundance, isoform expression patterns, protein interactions and post-translational modifications [1]. This makes proteomic datasets a fundamental part of modern systems biology and network analysis of biological processes [2], as well as the discovery of disease biomarkers [3]. Mass spectrometry is now the method of choice for swift and reliable protein identification in almost all high-throughput biochemical surveys of biological or pathological processes [4-6]. However, the routine proteomic investigation of crude tissue extracts can be complicated by sample complexity. Physicochemical properties of individual protein isoforms

can differ considerably within a given proteome [7]. Thus, the identification and characterization of thousands of proteins by standard separation techniques, such as liquid chromatography [8] or gel electrophoresis [9], often only achieves the coverage of the near-to-complete tissue proteome. Differences in size, charge, hydrophobicity and density seriously hamper the comprehensive identification of all protein species in a dynamic cellular system.

Organelle proteomics attempts to overcome these technical limitations by reducing sample complexity [10] using sophisticated pre-fractionation steps prior to proteomic analysis [11-13]. The analysis of membrane-associated proteins is especially challenging in large-scale proteomic studies [14-16] and is mostly due to the limited solubility, dy-

*Corresponding author: Dr. Kay Ohlendieck, Professor and Chair, Department of Biology, National University of Ireland, Maynooth, Co. Kildare, Ireland. Phone number: (353) (1) 708-3842; Fax number: (353) (1) 708-3845; E-mail address: kay.ohlendieck@nuim.ie

dynamic properties and relative low abundance of integral proteins in most biological systems [17-19]. Skeletal muscles contain a diverse range of protein constituents making the application of organelle proteomics an essential part of basic and applied myology. Voluntary contractile fibres exhibit a large dynamic range of proteins, a unique set of extremely high-molecular-mass proteins, numerous supramolecular membrane complexes and an extremely diverse isoform expression pattern of proteins involved in excitation-contraction coupling, contractile functions, muscle relaxation, ion homeostasis and metabolic pathways [20]. This review discusses the impact of recent subproteomic studies of normal, adapting and pathological skeletal muscle.

2. Defining the skeletal muscle proteome

Extensive mass spectrometry-based cataloging of the skeletal muscle proteome has resulted in the identification of thousands of fibre-associated protein isoforms, as well as the degree of crucial post-translational modifications, such as glycosylation, phosphorylation, tyrosine nitration and carbonylation [21]. Liquid chromatography, one-dimensional gels or two-dimensional gels using isoelectric focusing in the first dimension and slab gel electrophoresis in the second dimension are routinely employed to separate muscle proteins. The majority of fibre-associated proteins identified by proteomics belong to the class of contractile proteins of the thick and thin filaments (myosin light chains, myosin heavy chains, actins), regulatory proteins of the contractile apparatus (troponins and tropomyosins), enzymes of anaerobic metabolism (glycolytic enzymes), enzymes of oxidative metabolism (mitochondrial proteins), metabolic transportation (fatty acid binding proteins, albumin, myoglobin), as well as proteins responsible for detoxification and the cellular stress response (heat shock proteins). These studies included various muscles from humans [22] and animals that play a key role in biomedical research [23], as well as muscle samples crucial for the meat industry [24]. The proteomic comparison of fast versus slow muscles has revealed clear differences in protein expression patterns, showing altered levels and isoform distribution of key regulatory, functional and structural proteins belonging to the contractile apparatus, excitation-contraction coupling, ion homeostasis, cell signaling, stress response and muscle metabolism [25-27].

The proteomic characterization of human *vastus lateralis* muscle revealed that mitochondrial proteins accounted for 22% of the accessible skeletal muscle proteome [28]. This emphasizes the importance of mitochondria for muscle metabolism and shows that global proteomic studies can make an important contribution to our overall understanding of muscle physiology and bioenergetics. It also implies that organelle proteomics will play a key role in future studies of the skeletal muscle proteome during physiological adaptations or pathological insults. Thus, proteomic cataloguing studies are crucial for the establishment of comprehensive protein databanks for skeletal muscles in health and disease.

Besides the establishment of extensive proteomic maps of skeletal muscle based on gel electrophoretic studies, the application of shotgun proteomics to the biochemical characterization of human skeletal muscle has identified a large number of muscle-associated proteins. The proteomic analysis of biopsy specimens from 31 individuals has identified more than 2,000 protein species [29]. These databanks can now be employed as references for large-scale comparative studies of skeletal muscle samples. Whole tissue proteomics has investigated the effects of physical exercise, weight loss, muscular atrophy, muscle growth and fibre transitions, as well as the pathological impact of nerve damage, diabetes, sepsis, hypoxia, muscular dystrophy, inclusion body myositis, myotonia, age-related muscle wasting and motor neuron disease, as critically examined in several recent reviews [21, 22, 30-32].

3. Subproteomics of skeletal muscle tissues

Since all current large-scale biochemical separation methods have technical limitations with respect to properly separating all components within complex mixtures of heterogeneous proteins, routine proteomic studies of crude extracts do not usually cover the entire protein constellation of a given biological sample [33]. Although body liquids, such as saliva, urine or plasma with their almost exclusively soluble components, present exceptions and have been cataloged in their entirety by standard proteomics [34-36], the mass spectrometric identification of all proteins present in biological tissues is technically more challenging [37]. The extensive dynamic expression range of proteins in complex tissues and the diversity in charge, size and post-translational modifications of soluble proteins versus membrane proteins makes a reduction in sample complexity a prerequisite for inclusive proteomic studies. The subproteomes of distinct fractions from skeletal muscle have been described for nuclei [38], mitochondria [39-43], the contractile apparatus [44], sarcolemma [45], sarcoplasmic reticulum [46] and the cytosol [38, 47, 48], as described in detail in the below subsections.

3.1 Subcellular fractionation of skeletal muscle tissues

Muscle organelle proteomics attempts the cataloging and characterization of discrete subcellular fractions or supramolecular protein assemblies from contractile tissues, employing pre-fractionation steps prior to final protein separation. This may include micro-dissection approaches, differential centrifugation schemes, density gradient centrifugation, affinity isolation methods based on ligand or antibody technology, the usage of narrow pH ranges during the isoelectric focusing of low-abundance proteins, offgel electrophoresis for investigating proteins with extreme pI -values, detergent phase extraction or differential detergent fractionation techniques for studying integral proteins, and immobilization and blotting methods using on-membrane digestion for identifying extremely high-molecular-mass or very hydro-

phobic proteins [7, 11-13, 37, 49]. Emerging separation techniques are free-flow electrophoresis and fluorescent-assisted organelle sorting [50], which have not yet been widely applied to skeletal muscle biology. The flowchart of Figure 1 summarizes routine pre-fractionating steps for the isolation of major organelles, membrane systems and functional units from skeletal muscle homogenates. This includes the preparation of the contractile apparatus, nuclei, mitochondria, sarcoplasmic reticulum, transverse tubules, sarcolemma, extracellular matrix, cytoskeleton and the cytosolic fraction.

Standardized protocols for the subcellular fractionation of skeletal muscle homogenates consist of repeated centrifugation steps at progressively higher speeds and longer centrifugation periods. Distinct fractions enriched in nuclei, the contractile apparatus, mitochondria and microsomes are routinely achieved by centrifugation for 10 minutes at 1,000g, 10 minutes at 10,000g, 20 minutes at 20,000g and 1 hour at 100,000g, respectively. After high-speed centrifugation, the final supernatant contains almost exclusively components from the cytosolic fraction. Since myosins constitute one of the most abundant classes of proteins in muscle homogenates, myosin heavy and light chains often cross-

contaminate subcellular fractions following differential centrifugation. However, trapped or adsorbed myosin molecules can be easily removed from membrane preparations by mild salt washes [51]. The heterogeneous content of the microsomal fraction can be further separated by density gradient centrifugation. If a sufficiently high g-force is employed, surface membranes can be separated from the highly abundant sarcoplasmic reticulum, which usually divides into fractions enriched in longitudinal tubules, terminal cisternae and triad junctions [51]. The sarcolemma and transverse tubules can be further separated by affinity purification methods, such as differential lectin agglutination [52].

Subcellular marker proteins for distinct fractions from skeletal muscle tissue are represented by laminin for the basal lamina, the Na⁺/K⁺-ATPase for the sarcolemma, dystrophin for the sub-sarcolemmal membrane cytoskeleton, vimentin for the cytoskeleton, the dihydropyridine receptor for transverse tubules, the ryanodine receptor Ca²⁺-release channel for triad junctions, SERCA-type Ca²⁺-ATPases for longitudinal tubules of the sarcoplasmic reticulum, calsequestrin for the terminal cisterna region, sarcalumenin for the lumen of the sarcoplasmic reticulum, galactosyl transfer-

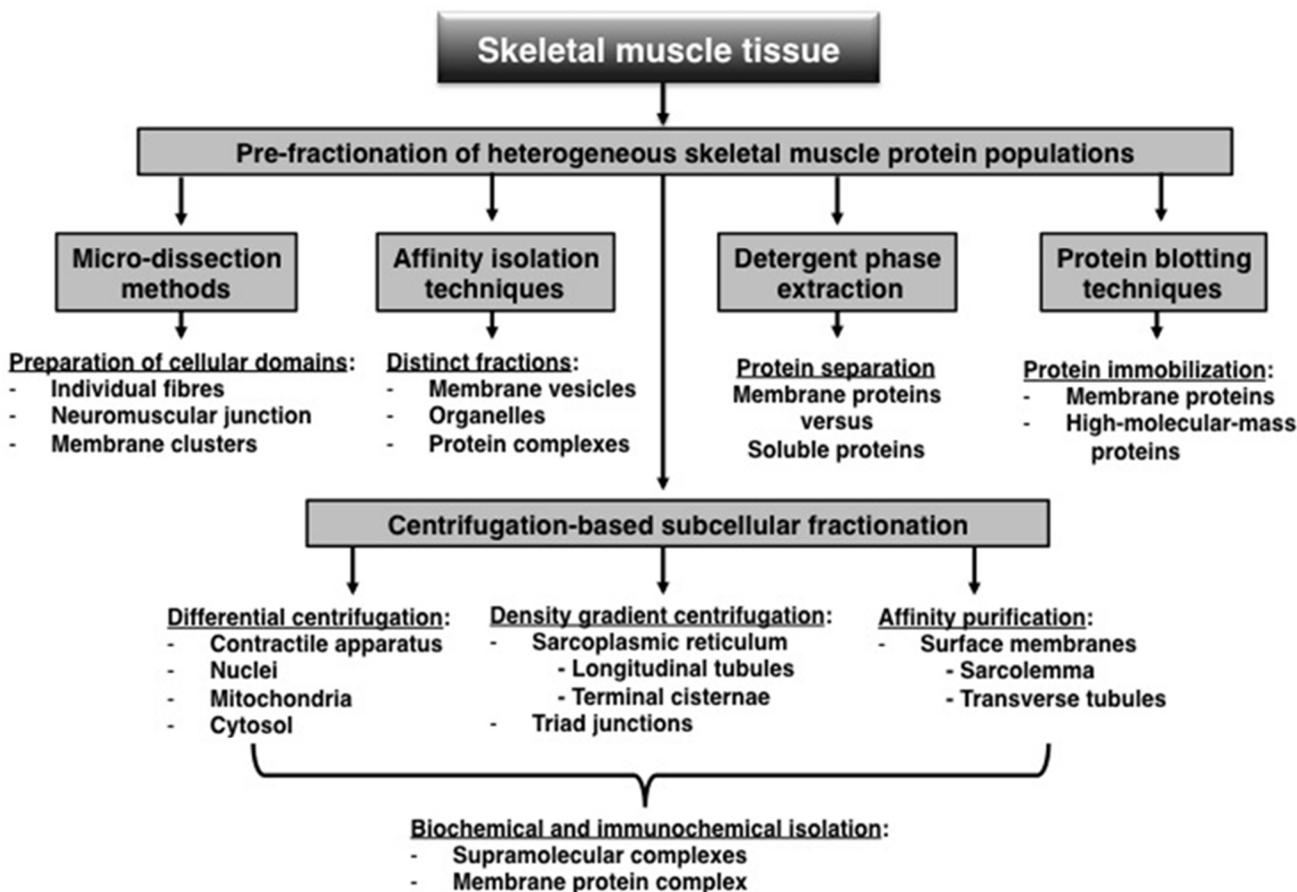


Figure 1. Overview of routine pre-fractionating techniques used for the isolation of major organelles, membrane systems and functional units from skeletal muscle. The flowchart summarizes the various experimental approaches employed to prepare fractions highly enriched in distinct organelles or large cellular structures, including micro-dissection methods, affinity isolation techniques, detergent extraction, protein blotting techniques and centrifugation-based subcellular fractionation protocols.

ase for the Golgi apparatus, histones for nuclei, ribosomal proteins for ribosomes, succinate dehydrogenase for mitochondria, acid phosphatase for lysosomes, catalase for peroxisomes, myosin heavy chains for the contractile apparatus and lactate dehydrogenase for the cytosol. For the subfractionation of intact organelles, the application of consecutive swelling, shrinking and sonication steps yields usually distinct membrane species. For example, mitochondrial subfractions can be identified by the enrichment of marker enzymes, such as succinate dehydrogenase of the inner membrane, glutathione transferase of contact sites and monoamino oxidase of the outer membrane.

3.2 Contractile apparatus

The distribution of myosin heavy chains, myosin light chains and fibre type-specific enzymes is indicative of the main fibre types I, IIa, IIx/d and IIb of skeletal muscles [53]. Since contractile proteins are extremely abundant in muscle fibres, comparative proteomic studies using crude tissue extracts routinely identify myosins, actins, troponins and tropomyosins by mass spectrometry [22]. Alterations in the abundance of specific contractile protein isoforms have been shown to occur during muscle adaptations, muscular atrophy and the natural aging process, as well as in a variety of neuromuscular diseases [21]. Previous biochemical studies have demonstrated that chronic low-frequency stimulation causes myosin light and heavy chains, as well as regulatory troponin subunits, to undergo a stepwise replacement from fast to slow isoforms [54]. Electrostimulation-induced muscle transformation has been confirmed by mass spectrometry-based proteomics and shown rapid fast-to-slow transitions due to enhanced neuromuscular activity [55, 56]. However, the detailed study of changed protein expression levels and/or isoform switching usually requires enriched preparations of the contractile apparatus.

Fractions enriched in myofibrillar proteins can be easily isolated from tissue homogenates by differential centrifugation [57, 58]. The proteomic characterization of isolated myofibrils from young adult versus senescent skeletal muscle has established a drastic increase in the abundance of slow myosin light chain MLC-2 during aging [44]. Phosphoproteomics could furthermore show that this contractile protein exhibits the highest change in phosphorylation during aging [44, 59]. The MLC-2 isoform of myosin light chain is an excellent indicator of fibre transformation in skeletal muscle [54] and its abundance changes and post-translational modifications establish this protein as an excellent candidate for being a suitable biomarker of sarcopenia of old age [60]. Comparative immunoblot analysis and immunofluorescence microscopy were used to verify proteomic data and confirmed a switch between fast and slow myosin heavy chains in aged muscle fibres [44]. Hence, these proteomic findings support the pathophysiological concept of an age-related shift to a slower-twitching fibre population that expresses predominantly slow isoforms within myosin hexamers [61].

In the long-term, the proteomic identification of altered concentration levels of individual muscle protein species may be useful in the biochemical evaluation of novel pharmacological or nutritional treatment options to counter-act age-related muscle wasting, as well as improved exercise regimes to prevent sarcopenia and frailty in the elderly.

3.3 Skeletal muscle mitochondria

Global changes in mitochondrial proteins are intensively studied [62, 63]. A recent review by Silvestri et al. [64] has described in detail the importance of various analytical approaches, including two-dimensional gel electrophoresis, blue native polyacrylamide gel electrophoresis, fluorescence difference in-gel electrophoresis, shotgun proteomics and stable isotope labeling analysis, mitochondrial protein arrays and mitochondrial protein databanks. The mammalian mitochondrial proteome consists of over 1,000 different proteins [65-67], which are distributed throughout the inner matrix region, a heavily folded inner membrane structure, a smooth outer membrane and contact sites between the two membranes, as well as an inter-membrane space that is continuous with the cytosol [68, 69]. While a small number of mitochondrial proteins are encoded by a unique mitochondrial genome [70], a large population of proteins is imported into mitochondria via sophisticated transport mechanisms [71, 72]. Mitochondria represent one of the most diverse class of organelles in the body with respect to multiple metabolic functions and manifold involvements in regulatory processes. Having originated as endosymbionts of a primordial eukaryote, highly evolved mitochondria play a key metabolic role as the primary site for energy generation via oxidative phosphorylation [68]. They are furthermore involved in the integration of intermediary metabolism, protein transport, the production of heme and iron-sulfur clusters, cell cycle progression, calcium ion signaling and the regulation of apoptosis [69]. The proteome of mitochondria exhibits considerable tissue heterogeneity [39, 40] and the protein constellation of skeletal muscle mitochondria can be distinguished by their subsarcolemmal versus intermyofibrillar location [73]. Interestingly, proteomics has shown that the mitochondrial oxidative phosphorylation system is associated with sarcolemmal lipid rafts during myogenesis [74] and that muscle mitochondria are closely coupled to lipid droplets probably promoting metabolic channeling [75].

Proteomic profiling of human *vastus lateralis* muscle has identified 823 distinct mitochondrial proteins [41], making mitochondria one of the most extensively catalogued organelle from voluntary contractile tissue [29]. Post-translational modifications have also been extensively studied in isolated muscle mitochondria by mass spectrometry [76, 77]. Since altered numbers of mitochondria, functional alterations within this crucial organelle and/or changed expression levels within the mitochondrial proteome play a central role in various disorders [67, 78] and especially during the natural aging process of skeletal muscles [79-81],

numerous subproteomic studies have targeted mitochondria and their involvement in senescence of the neuromuscular system [82]. In agreement with comparative studies of total tissue extracts from adult versus aged muscle [21, 22], organelle proteomics identified a large number of changed mitochondrial proteins during aging [42, 43]. Proteomic profiling in combination with Blue Native PAGE analysis, which has been widely applied to defining the mitochondrial protein population [83], revealed interesting age-related changes in the abundance of mitochondrial complexes involved in oxidative phosphorylation [43]. The mitochondria-enriched fraction from senescent rat muscle was shown to exhibit drastically increased levels of NADH dehydrogenase, succinate dehydrogenase, creatine kinase and ubiquinol cytochrome-c reductase [42]. In addition, the offgel electrophoretic analysis of basic proteins could also show increased levels of mitochondrial creatine kinase and ubiquinol cytochrome-c reductase in aged muscle [84]. These findings agree with the pathobiochemical concept of a drastic metabolic shift during fibre aging [61]. Skeletal muscle tissues appear to perform a higher degree of aerobic-oxidative metabolism in a slower-twitching fibre population during age-dependent fibre degeneration [60]. However, this process is probably not an adaptive mechanism during aging, but rather a pathological consequence of the selective denervation and preferential loss of type II fibres in senescent muscles [85]. The progressive loss of muscle mass and function in the elderly has now been termed sarcopenia and is believed to play a major pathophysiological role in the overall functional decline of the senescent organism [86-88].

Regular physical activity and a healthy lifestyle play a crucial role in preventative medicine, including the avoidance of an early onset of certain forms of diabetes, obesity or sarcopenia of old age. Thus, studying the effects of endurance exercise versus disuse atrophy by proteomics can give crucial molecular and cellular insights into global mechanisms of muscle adaptation [31]. Changed functional demands have profound effects on protein expression levels and proteomic studies have shown that on the one hand chronic neuromuscular activity triggers fast-to-slow transitions [56, 89] and on the other hand muscular atrophy is clearly associated with slow-to-fast transformation processes [90-92]. The fact that endurance training triggers extensive modifications of the mitochondrial protein constellation was recently confirmed by organelle proteomics. Egan et al. [93] have shown that a large number of mitochondrial markers are increased in human *vastus lateralis* muscle following endurance exercise, including creatine kinase, ATP synthase and malate dehydrogenase. These proteomic findings support the idea of skeletal muscle tissue being highly plastic in its physiological response and that the neuromuscular system can quickly adapt to new physical challenges.

In contrast to physical training, muscular disuse and obesity has been linked to a higher susceptibility to type 2 diabetes. Diabetes mellitus does not only affect the heart in a major way, but also triggers contractile weakness and metabolic

disturbances in skeletal muscles [94]. Muscle proteomics has shown major changes in the protein profile of patients with insulin resistance [95], whereby an oxidative-to-glycolytic shift appears to occur in diabetic fibres [96]. The proteomic profiling of isolated mitochondria from diabetic skeletal muscle has shown a generally perturbed protein expression pattern with a reduction in NADH dehydrogenase, cytochrome b-c1 complex and isocitrate dehydrogenase. Isoforms of ATP synthase and pyruvate dehydrogenase showed differential changes in their expression levels [97]. Mitochondrial abnormalities probably cause a diabetes-related impairment of oxidative phosphorylation and may thus be crucial for metabolic disturbances and the development of insulin resistance [32].

3.4 Surface membranes

The highly complex arrangement of the sarcolemma and its membranous invaginations, the transverse tubular system, is involved in a plethora of cellular functions in skeletal muscle fibres. This includes synaptic transmission, propagation of action potentials, excitation-contraction coupling, the essential stabilization of the fibre periphery during contraction-relaxation cycles, indispensable nutrient uptake, metabolite transportation, the maintenance of cell signaling systems, the provision of natural repair mechanisms following membrane rupturing, the regulation of ion homeostasis and general structural support for the preservation of fibre integrity [98]. Since the sarcolemma and transverse tubules are relatively low in abundance as compared to the intricate membrane system of the sarcoplasmic reticulum, the biochemical isolation of distinct surface membrane vesicles from skeletal muscle homogenates without cross-contaminations is a difficult task. However, an important prerequisite of meaningful muscle organelle proteomics is the biochemical purity of the starting material for protein analysis [7, 37]. Hence impurities due to non-specific protein adsorption and/or cross-contaminations due to membrane entrapments during isolation procedure have to be kept to a minimum. Ideally suited for such applications are affinity isolation methods that exploit distinct cellular, immunological or physicochemical properties of subcellular fractions and their constituents.

Affinity isolation of plasmalemma vesicles in combination with one-dimensional gradient gel electrophoresis and on-membrane digestion has been used to determine the protein composition of highly purified sarcolemma vesicles from rabbit skeletal muscle [45]. This novel approach has overcome several technical limitations of standard gel electrophoresis-based proteomics that routinely employs two-dimensional gel electrophoresis and subsequent in-gel digestion for mass spectrometric studies. Two-dimensional gel electrophoresis is an excellent method for the large-scale separation of urea-soluble and abundant proteins that fall into a molecular mass range of approximately 10 to 200 kDa. However, proteins with a low copy number, very high mo-

lecular mass, mostly hydrophobic properties, extensive post-translational modifications and/or extreme *pI*-values are poorly represented by this electrophoretic technique. Thus, for studying integral proteins of high molecular mass it is advantageous to pre-fractionate tissue samples to reduce sample complexity and then use large one-dimensional gradient gels for protein separation. In addition, since high-throughput proteomic surveys using in-gel digestion procedures are sometimes complicated by an inefficient trypsination of certain protein species, on-membrane digestion with superior protein sequence coverage can be used as an alternative [99-101].

When studying highly enriched sarcolemma vesicles from skeletal muscle, the combination of lectin affinity purification, one-dimensional gradient gel electrophoresis and on-membrane digestion has resulted in the comprehensive cataloguing of major protein bands representing the plasmalemma [45]. Mass spectrometric screening of sarcolemma-associated proteins identified a number of very large muscle components, including the membrane cytoskeletal element dystrophin of 427 kDa. The Dp427 isoform, the absence of which causes x-linked muscular dystrophy, had not previously been identified in numerous proteomic studies of normal versus dystrophic muscles due to its very high molecular mass [30]. Hence, affinity organelle isolation and on-membrane digestion were shown to be highly appropriate for the proteomic identification of high-molecular-mass proteins that would otherwise not be properly recognized by standard two-dimensional gel electrophoresis.

3.5 Sarcoplasmic reticulum

The specialized endoplasmic reticulum of skeletal muscles, the sarcoplasmic reticulum, is a highly abundant organellar structure and acts as a crucial physiological controller of Ca^{2+} -cycling throughout the muscle interior [102]. The sarcoplasmic reticulum thus plays a key role in the regulation of the excitation-contraction-relaxation cycle of muscle fibres [103]. Since the contractile status of muscles is determined by cytosolic Ca^{2+} -levels, the spatiotemporal organization of Ca^{2+} -release versus energy-dependent Ca^{2+} -uptake by the sarcoplasmic reticulum regulates excitation-contraction coupling and fibre relaxation. A large number of Ca^{2+} -channels, Ca^{2+} -ATPases, Ca^{2+} -exchangers and Ca^{2+} -binding proteins are involved in the highly complex maintenance of Ca^{2+} -homeostasis in skeletal muscle tissues [104]. Transient opening of the nicotinic acetylcholine receptor by neurotransmitter binding triggers the massive influx of Na^{+} -ions into muscle fibres at the neuromuscular junction, which in turn causes the propagation of an action potential along the sarcolemma via voltage-dependent Na^{+} -channel activation. Within the transverse tubular membrane, the voltage-sensing α_{1S} -subunit of the dihydropyridine receptor interacts with the ryanodine receptor complex by direct physical means through its II-III loop domain and triggers the opening of the Ca^{2+} -release channel at the triad junction [105]. Passive

efflux of Ca^{2+} -ions along a steep gradient raises the level of this second messenger in the cytosol and initiates muscle contraction by binding to the troponin TnC subunit. Fibre relaxation is caused by the active re-uptake of Ca^{2+} -ions into the lumen of the sarcoplasmic reticulum by slow SERCA2 and fast SERCA1 Ca^{2+} -ATPase complexes. Ion shuttling within the organelle is provided by sarcalumenin and ion storage and channeling to the ryanodine receptor is mediated by the high-capacity Ca^{2+} -binding protein calsequestrin [106].

As already outlined in above section on sarcolemma proteomics, standard gel electrophoresis-based proteomics does not cover all protein species present in crude extracts or sub-cellular fractions. This is especially problematic in the case of very large integral proteins of which the sarcoplasmic reticulum contains many examples, such as the ryanodine receptor Ca^{2+} -release channel monomer of 565 kDa [107]. In order to be able to identify this hydrophobic protein of the sarcoplasmic reticulum by mass spectrometry, a combination of sub-cellular fractionation, gradient gel electrophoresis, protein blotting and on-membrane digestion had to be employed [46]. The proteomic survey of the sarcoplasmic reticulum revealed the presence of 31 distinct protein species, including all major Ca^{2+} -regulatory proteins involved in the excitation-contraction-relaxation cycle, such as Ca^{2+} -ATPase, calsequestrin, sarcalumenin and the Ca^{2+} -release channel. Previous ultrastructural studies had localized glycolytic enzymes on sarcoplasmic reticulum vesicles [108]. The biochemical concept of a close physical coupling between the energy-dependent sarcoplasmic reticulum and the ATP-producing glycolytic pathway was confirmed by proteomics. Mass spectrometry clearly identified the presence of the glycolytic enzymes phosphofruktokinase and aldolase in the purified sarcoplasmic reticulum [46]. Thus, organelle proteomics using on-membrane digestion methodology is an excellent way for studying high-molecular-mass proteins and hydrophobic proteins.

The cationic carbocyanine dye 'Stains-all' labels most gel-bound proteins with a light pinkish colour, but stains Ca^{2+} -binding proteins with a characteristic dark purple colour whereby dye-protein complexes absorb maximally at 615 nm [109]. This property of 'Stains-all' dye, in combination with mass spectrometry and immunoblotting, was exploited in a subproteomic study of the fate of Ca^{2+} -binding proteins in dystrophic muscle [110]. Duchenne muscular dystrophy is an x-linked inherited muscle wasting disease and due to primary abnormalities in the membrane cytoskeletal protein dystrophin [111]. Micro-rupturing of the dystrophin-deficient sarcolemma and inefficient repair mechanisms are believed to cause irregular ion handling, which is probably a key pathophysiological factor that renders skeletal muscle fibres more susceptible to necrosis [112]. Comparative subproteomics demonstrated that the reduced Ca^{2+} -buffering capacity of the sarcoplasmic reticulum from dystrophic fibres is caused by drastically decreased levels of the main luminal Ca^{2+} -reservoir protein calsequestrin [110]. Proteomic

analyses have also established that the luminal Ca^{2+} -shuttle protein sarcalumenin is reduced in dystrophin-deficient muscle [113]. Hence, a reduction in essential luminal Ca^{2+} -binding proteins probably plays a key role in the molecular pathogenesis of x-linked muscular dystrophy. With respect to potential therapeutic implications, it is important to mention that the proteomic evaluation of experimental exon-skipping treatment has shown a reversal of calsequestrin reduction in treated fibres [114]. There appears to be a direct connection between the re-establishment of the subsarcolemmal membrane cytoskeleton and the Ca^{2+} -handling apparatus of the sarcoplasmic reticulum in muscle following reversal of the primary defect in dystrophinopathy [30].

3.6 Muscle cytosol

The sarcoplasm provides the aqueous environment within a muscle fibre and supports and surrounds organelles with its semi-fluid material containing water, anions, cations, organic molecules, nutrients and numerous metabolic and signaling enzymes. The cytoplasm functions as a polar solvent for soluble cellular constituents and facilitates the transportation of essential metabolites and messenger molecules within the contractile fibre. A key anaerobic pathway for energy production and metabolic integration is glycolysis, which occurs almost exclusively in the cytosol of muscle fibres. The enzyme hexokinase is the only exception and its activity levels are regulated by binding to the mitochondrial outer membrane for metabolic channeling purposes. Subproteomic studies have confirmed the high abundance and central metabolic position of glycolytic enzymes in skeletal muscle tissues. Proteomic analysis has clearly established that the most abundant protein species present in the diffusible fraction of the skeletal muscle proteome are the 10 glycolytic enzymes [47]. The high density and solubility of the enzymes that mediate the core glycolytic pathway makes them ideal biomarker candidates to be evaluated by standard biochemical separation methods and mass spectrometry, which has been covered in an extensive review [115].

A subproteomic study by Vitorino et al. [38] of the nuclear, mitochondrial and cytosolic fractions from fast versus slow muscles has established a considerable number of cytosol-associated proteins. Although the application of conventional subcellular fractionation methods results in a considerable degree of cross-contaminations between individual fractions, these kinds of proteomic maps may become useful in future comparative studies for evaluating the effects of physiological adaptations or pathological insults on the muscle cytosol. Previously, cytosolic protein changes were studied in the mdx animal model of Duchenne muscular dystrophy [116, 117] and atrophying mouse skeletal muscle [48] by subproteomics. The proteomic comparison of 3-month old dystrophic mdx muscles versus age-matched normal controls revealed the differential expression of approximately 40 proteins from the cytosolic fraction [116, 117]. Abnormal levels of adenylate kinase isoform AK1 and creatine kinase

indicate disturbed regulation of nucleotide ratios and energy metabolism in dystrophin-deficient muscle tissue [116] and these altered protein levels were also shown to exist in the severely dystrophic mdx diaphragm muscle [118, 119].

3.7 Skeletal muscle secretome

The proteins secreted by cells and tissues have a great potential to be exploited as disease- and stage-specific biomarkers of a variety of pathological conditions [120]. The entirety of molecules secreted from living muscle fibres are referred to as the muscle secretome, whereby released proteins are involved in myogenesis, cellular signaling, cell-cell communication, proliferation and cell migration [121]. This makes the cataloguing of the fibre secretome an important aspect of studying autocrine and paracrine signaling mechanisms of the neuromuscular system. For proteomic studies of the muscle secretome, the collection of the released protein fraction under controlled conditions is challenging and often hampered by cross-contamination and/or interference by experimentally uncontrollable cellular events other than true protein secretion. Besides these technical problems, initial proteomic studies focusing on the muscle secretome have identified interesting groups of extracellular factors that may be involved in the triggering and regulation of intracellular events involved in development, muscle repair and fibre adaptations [122-126]. Secretome proteomics has been especially applied to understanding the role of external factors during myogenesis, including myoblast proliferation, myoblast differentiation and myotube formation [121].

The proteomic survey of cultured C2C12 muscle cells that were grown in a serum-free medium resulted in the identification of a large number of secretory proteins involved in extracellular matrix remodeling, cellular proliferation, migration, and cellular signaling [122]. Quantitative proteomic analysis of the dynamic muscle secretome at different time points of myoblast differentiation identified over 600 released proteins, including cytokines, growth factors and metallo-peptidases [123]. The proteomic identification of changes in the human myotube secretome due to exposure to insulin or tumor necrosis factor $\text{TNF-}\alpha$ revealed several new mediator candidates that are being secreted at lower or higher levels following insulin stimulation or during insulin resistance [125, 126]. This indicates that skeletal muscle is a prominent secretory organ that produces a considerable amount of extracellular factors that mediate signaling to the highly complex contractile system and surrounding tissues.

3.8 Proteomics of select protein populations and protein complexes

Comparative biochemical studies are ideally performed with crude total tissue extracts in order to avoid potential artifacts due to extensive subcellular fractionation procedures. Extensive tissue manipulation and numerous preparative steps in the enrichment of a particular organelle or sub-

cellular fraction may cause uncontrolled protein release, protein adsorption, protein entrapment or non-specific protein loss. Preparation of tissue homogenates and the subsequent isolation of membrane fractions often result in the production of mixed vesicle populations, including membrane sheets, leaky vesicles, inside-out vesicles, right-side-out vesicles and smaller vesicles entrapped in larger membrane assemblies. Thus, in comparative organelle studies these technical complications have to be taken into account and may seriously limit the proper evaluation of protein levels in physiologically challenged or pathologically damaged muscle preparations. In order to reach a high degree of validity, subproteomic studies have to use highly purified samples with a minimum of impurities from other cellular compartments. An alternative to organelle proteomics is the biochemical isolation of distinct protein populations or the preparation of the entire protein complement from a given tissue sample.

Filter-aided sample preparation is a method that exploits a combination of detergent solubilization of the entire protein constellation of a cell, retention and concentration of the sample on an ultra-filtration device, detergent removal, exchange with urea buffer, chemical modifications and finally controlled protein digestion for mass spectrometric analysis [127]. This procedure covers theoretically the entire proteome of a given sample. If a proteomic analysis focuses on soluble versus hydrophobic proteins, detergent phase extraction would be the method of choice. This separation technique exploits the principle of temperature-dependent phase extraction with the detergent Triton X-114. Phase separation occurs at temperatures above 22°C using Triton X-114. A detergent phase with predominantly hydrophobic proteins and an aqueous phase enriched in hydrophilic proteins occurs reproducibly at an experimental temperature setting of 37°C [128]. As can be seen with most separation methods, detergent phase transition approaches also cause a certain degree of cross-contamination between integral and soluble protein species. The fluorescence difference in-gel electrophoretic analysis of aging skeletal muscle proteins using nonionic detergent phase extraction revealed alterations in a large number of muscle-associated proteins involved in energy metabolism, metabolic transportation, regulatory processes, the cellular stress response, detoxification mechanisms and fibre contraction [129].

Skeletal muscle tissues contain several extremely large membrane-associated protein complexes, such as the nicotinic acetylcholine receptor of the post-synaptic membrane folds, the ryanodine receptor Ca^{2+} -release channel of the triad junctions and the dystrophin-glycoprotein complex that encompasses the membrane cytoskeleton, the sarcolemma and the extracellular matrix. Proteomics has been used to characterize the supramolecular dystrophin-associated complex and confirmed the composition of this protein assembly by mass spectrometry [45, 130]. Immunoprecipitation was used to isolate sarcolemmal β -dystroglycan and its tightly associated members of the dystrophin-glycoprotein complex

from detergent-solubilized skeletal muscle [130]. Another proteomic approach used biochemically purified dystrophin and its associated glycoproteins and identified individual components by mass spectrometry following gradient gel electrophoresis and on-membrane digestion [45]. Proteins from digitonin-solubilized muscle membranes were separated by ion exchange chromatography, lectin binding and sucrose gradient centrifugation. The mass spectrometric analysis clearly established a tight linkage between dystrophin of 427 kDa and integral glycoproteins of the sarcolemma [45]. This demonstrates that mass spectrometry-based proteomics can be successfully applied to the identification and characterization of relatively low-abundance and membrane-associated muscle protein complexes. These new isolation and detection methods can be extremely useful for detailed future studies into the pathobiochemical role of the dystrophin complex in muscular dystrophy.

3.9 Proteomic markers of muscle organelles

Skeletal muscle proteomics has both confirmed the usefulness of established subcellular markers in high-throughput studies and identified novel candidates for the characterization of organelles and membrane systems in large-scale biochemical surveys. The diagram of Figure 2 gives an overview of select subcellular markers of skeletal muscle fibres. As listed in Table 1, this includes proteins associated with the contractile myosin-containing filaments (myosin heavy chains, myosin light chains), contractile actin-containing filaments (various actin isoforms), regulatory complexes of the contractile apparatus (tropomyosin TM, troponin subunits TnT, TnI, TnC), the sarcolemma membrane (dysferlin, Na^+/K^+ -ATPase, β -dystroglycan), the sub-sarcolemmal membrane cytoskeleton (full-length dystrophin isoform Dp427), the cytoskeleton (vimentin, desmin), the extracellular matrix (laminin, collagen), the transverse tubular membrane system (voltage-sensing dihydropyridine receptor complex), the extracellular space (albumin), the longitudinal tubules membranes of the sarcoplasmic reticulum (SERCA-type Ca^{2+} -ATPases), the terminal cisternae region of the sarcoplasmic reticulum (Ca^{2+} -binding protein calsequestrin), the lumen of the sarcoplasmic reticulum (Ca^{2+} -shuttle protein sarcalumenin), the triad junction membrane (ryanodine receptor RyR1 Ca^{2+} -release channel complex), the outer membrane of mitochondria (VDAC porin), the inner membrane of mitochondria (Complex I, NADH dehydrogenase; Complex II, succinate dehydrogenase; Complex III, cytochrome b-c complex; Complex IV, cytochrome c oxidase; and Complex V, ATP synthase), and the mitochondrial matrix (isocitrate dehydrogenase), as well as cytosolic compartments that provide metabolite transportation (myoglobin, fatty acid binding protein), the cellular stress response (molecular chaperones) and metabolic pathways (glycolytic enzymes).

Table 1. List of marker proteins representative of distinct subcellular fractions and pathways that are routinely identified by muscle proteomics

Subcellular muscle fraction	Muscle-associated marker proteins
Contractile apparatus	Myosin heavy chains: MHC I, IIa, IIx, IIb
Myosin filament	Myosin light chains: MLC 1f, 2f, 3f, 1s, 2s
Contractile apparatus	Actins: various α isoforms
Actin filament	
Contractile apparatus	Tropomyosins: TM as, af, b
Tropomyosin filament	
Contractile apparatus	Troponin subunits: TnT (1f-4f), TnI (f, s), TnC (f, s)
Troponin complex	
Sarcolemma membrane	Dysferlin, Na ⁺ /K ⁺ -ATPase, β -dystroglycan
Sub-sarcolemmal membrane	Dystrophin (full-length isoform Dp427)
cytoskeleton	
Cytoskeleton	Vimentin, desmin
Transverse tubular membrane	Dihydropyridine receptor complex
Extracellular space	Albumin
Extracellular matrix	Collagen, laminin
Sarcoplasmic reticulum	SERCA-type Ca ²⁺ -ATPases (SERCA1, SERCA2)
longitudinal tubules membrane	
Sarcoplasmic reticulum	Calsequestrin Ca ²⁺ -binding protein (CSQf, CSQs)
terminal cisternae region	
Sarcoplasmic reticulum lumen	Sarcolumenin Ca ²⁺ -binding protein (SAR)
Triad junction membrane	Ryanodine receptor Ca ²⁺ -release channel complex (Muscle isoform RyR1)
Mitochondria - outer membrane	VDAC porin (VDAC1 isoform)
Mitochondria - inner membrane	NADH dehydrogenase
Complex I	
Mitochondria - inner membrane	Succinate dehydrogenase
Complex II	
Mitochondria - inner membrane	Cytochrome b-c complex
Complex III	
Mitochondria - inner membrane	Cytochrome c oxidase
Complex IV	
Mitochondria - inner membrane	ATP synthase
Complex V	
Mitochondria - matrix	Isocitrate dehydrogenase
Cytosol - metabolite transportation	Myoglobin, fatty acid binding protein FABP3
Cytosol - molecular chaperones	Small heat shock proteins (α B-crystallin, cvHsp)
Cytosol - glycolytic particle	Glycolytic enzymes

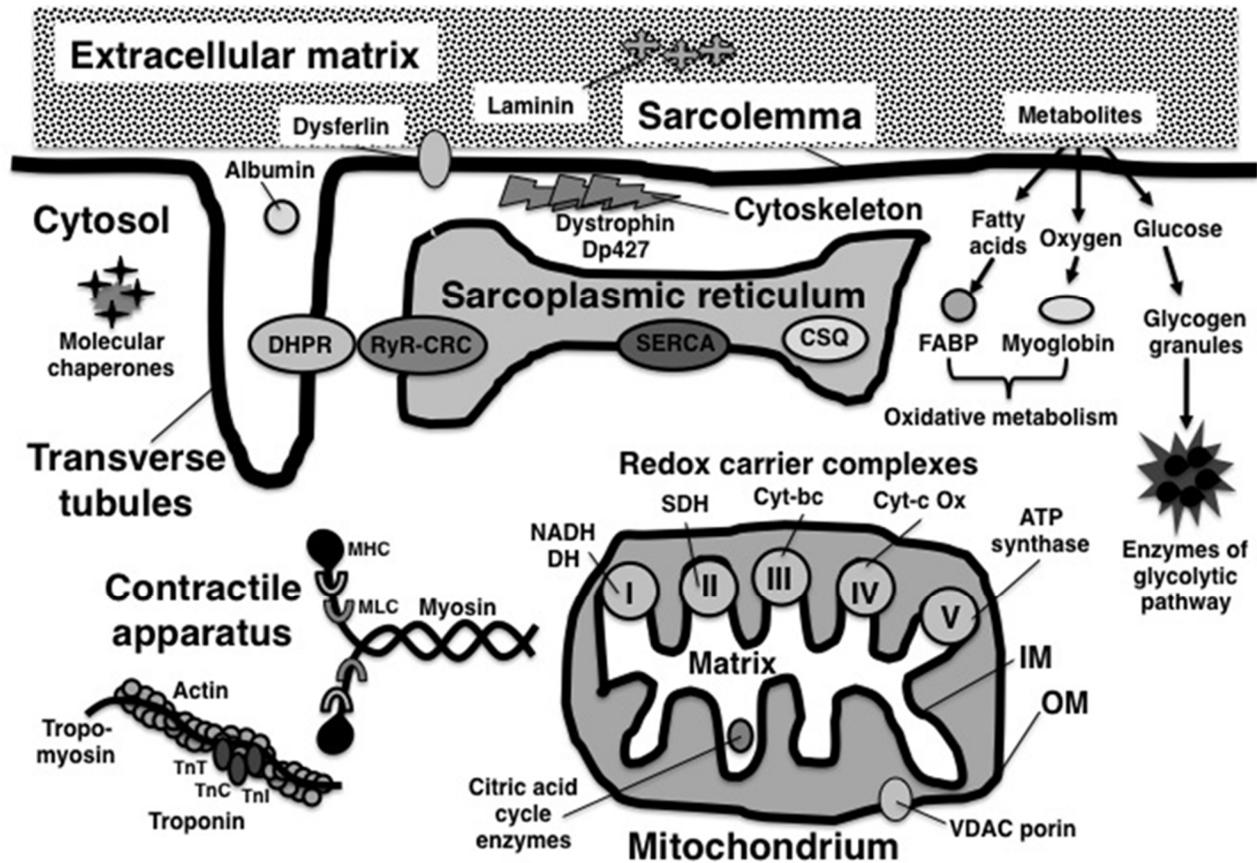


Figure 2. Overview of proteomic markers representative of distinct subcellular fractions from skeletal muscle. The diagram summarizes the subcellular position of established and frequently used marker proteins of the extracellular matrix, extracellular space, sarcolemma, transverse tubules, cytoskeleton, sarcoplasmic reticulum, contractile apparatus, mitochondria and cytosol.

4. Future perspectives

The determination of global cellular mechanisms and proteomic biomarker identification are now at the forefront of many large-scale biochemical surveys of normal, adapting and pathological skeletal muscles. However, the dynamic expression range of muscle proteins and the extremely diverse biochemical properties of individual proteins with respect to size, charge, hydrophobicity and extent of post-translational modifications seriously complicate the comparative proteomic analysis of whole tissues. Thus, in the foreseeable future the usage of subproteomics for studying isolated organelles and distinct cellular fractions will be a necessity for covering the majority of constituents in a tissue proteome. In order to improve the analytical impact of organelle proteomics, it will be essential to utilize more precise mass spectrometric approaches even for the identification of proteins of very low abundance, but also develop superior subcellular fractionation techniques that exhibit a minimum degree of cross-contamination and employ wide-ranging separation schemes that coalesce the technical advantages of gel electrophoresis, liquid chromatography and optimum protein digestion for identifying a maximum number of proteins. The miniaturization of isolation methods, especially in the preparation of individual muscle fibres with an intact

morphology and the microscopical dissection of cellular domains, promises to have a considerable impact on organelle proteomics. Once more comprehensive databanks of muscle organelle proteomes have been established, it will be crucial to use sophisticated bioinformatics to better integrate findings from genomics, transcriptomics, proteomics, metabolomics and cytomics for the systems biological evaluation of fundamental aspects of muscle biology and neuromuscular pathology.

Acknowledgements

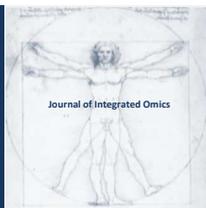
The author would like to thank the Irish Health Research Board, the Irish Higher Education Authority (PRTL15 Bio-AT programme), Science Foundation Ireland and Muscular Dystrophy Ireland for continued funding of our proteomics research projects focusing on skeletal muscle tissues in health and disease.

References

1. T.E. Angel, U.K. Aryal, S.M. Hengel, E.S. Baker, R.T. Kelly, E.W. Robinson, R.D. Smith, *Chem. Soc. Rev.* 41 (2012) 3912-3928.
2. A. Bensimon, A.J. Heck, R. Aebersold, *Annu. Rev. Biochem.*

- 81 (2012) 379-405.
3. D. Calligaris, C. Villard, D. Lafitte, J. Proteomics 74 (2011) 920-934.
 4. X. Han, A. Aslanian, J.R. Yates 3rd, Curr. Opin. Chem. Biol. 12 (2008) 483-490.
 5. I.A. Brewis, P. Brennan, Adv. Protein Chem. Struct. Biol. 80 (2010) 1-44.
 6. B.T. Chait, Annu. Rev. Biochem. 80 (2011) 239-246.
 7. J.R. Yates 3rd, A. Gilchrist, K.E. Howell, J.J. Bergeron, Nat. Rev. Mol. Cell Biol. 6 (2005) 702-714.
 8. F. Xie, T. Liu, W.J. Qian, V.A. Petyuk, R.D. Smith, J. Biol. Chem. 286 (2011) 25443-25449.
 9. V.J. Gauci, E.P. Wright, J.R. Coorsen, J. Chem. Biol. 4 (2011) 3-29.
 10. A.F. Altelaar, A.J. Heck, Curr. Opin. Chem. Biol. 16 (2012) 206-213.
 11. D.J. Gauthier, C. Lazure, Expert Rev. Proteomics 5 (2008) 603-617.
 12. U. Michelsen, J. von Hagen, Methods Enzymol. 463 (2009) 305-328.
 13. Y.H. Lee, H.T. Tan, M.C. Chung, Proteomics 10 (2010) 3935-3956.
 14. S. Tan, H.T. Tan, M.C. Chung, Proteomics 8 (2008) 3924-3932.
 15. Y.Z. Zheng, L.J. Foster, J. Proteomics 72 (2009) 12-22.
 16. P. Souda, C.M. Ryan, W.A. Cramer, J. Whitelegge, Methods 55 (2011) 330-336.
 17. A.O. Helbig, A.J. Heck, M. Slijper, J. Proteomics 73 (2010) 868-878.
 18. N.P. Barrera, C.V. Robinson, Annu. Rev. Biochem. 80 (2011) 247-271.
 19. J.N. Savas BD Stein, C.C. Wu, J.R. Yates 3rd, Trends Biochem. Sci. 36 (2011) 388-396.
 20. P. Doran, P. Donoghue, K. O'Connell, J. Gannon, K. Ohlendieck, Int. J. Mol. Med. 19 (2007) 547-564.
 21. K. Ohlendieck, Expert Rev. Proteomics 7 (2010) 283-296.
 22. C. Gelfi, M. Vasso, P. Cerretelli, J. Proteomics 74 (2011) 774-795.
 23. P. Doran, J. Gannon, K. O'Connell, K. Ohlendieck, Proteomics Clin. Appl. 1 (2007) 1169-1184.
 24. B. Picard, C. Berri, L. Lefaucheur, C. Molette, T. Sayd, C. Terlouw, Brief. Funct. Genomics 9 (2010) 259-278.
 25. N. Okumura, A. Hashida-Okumura, K. Kita, M. Matsubae, T. Matsubara, T. Takao, K. Nagai, Proteomics 5 (2005) 2896-2906.
 26. D. Capitanio, A. Vigano, E. Ricci, P. Cerretelli, R. Wait, C. Gelfi, Proteomics 5 (2005) 2577-2586.
 27. C. Gelfi, A. Vigano, S. De Palma, M. Ripamonti, S. Begum, P. Cerretelli, R. Wait, Proteomics 6 (2006) 321-340.
 28. K. Højlund, Z. Yi, H. Hwang, B. Bowen, N. Lefort, C.R. Flynn, P. Langlais, S.T. Weintraub, L.J. Mandarino, Mol. Cell. Proteomics 7 (2008) 257-267.
 29. K.C. Parker, R.J. Walsh, M. Salajegheh, A.A. Amato, B. Krastins, D.A. Sarracino, S.A. Greenberg, J. Proteome Res. 8 (2009) 3265-3277.
 30. C. Lewis, S. Carberry, K. Ohlendieck, J. Muscle Res. Cell. Motil. 30 (2009) 267-279.
 31. J.G. Burniston, E.P. Hoffman, Expert Rev. Proteomics 8 (2011) 361-377.
 32. K. Ohlendieck, J. Nutr. Metab. 2012 (2012) 893876.
 33. K. Ohlendieck, Skelet Muscle 1 (2011) 6.
 34. M. Prudent, J.D. Tissot, N. Lion, Expert Rev. Proteomics 8 (2011) 717-737.
 35. W. Mullen, Expert Rev. Proteomics 9 (2012) 375-387.
 36. W.L. Siqueira, C. Dawes, Proteomics Clin. Appl. 5 (2011) 575-579.
 37. A.J. Groen, K.S. Lilley, Expert Rev. Proteomics 7 (2010) 867-878.
 38. R. Vitorino, R. Ferreira, M. Neuparth, S. Guedes, J. Williams, K.B. Tomer, P.M. Domingues, H.J. Appell, J.A. Duarte, F.M. Amado, Anal. Biochem. 366 (2007) 156-169.
 39. F. Forner, L.J. Foster, S. Campanaro, G. Valle, M. Mann, Mol. Cell. Proteomics 5 (2006) 608-619.
 40. N.H. Reifschneider, S. Goto, H. Nakamoto, R. Takahashi, M. Sugawa, N.A. Dencher, F. Krause, J. Proteome Res. 5 (2006) 1117-1132.
 41. N. Lefort, Z. Yi, B. Bowen, B. Glancy, E.A. De Filippis, R. Mapes, H. Hwang, C.R. Flynn, W.T. Willis, A. Civitarese, K. Højlund, L. J. Mandarino, J. Proteomics 72 (2009) 1046-1060.
 42. K. O'Connell, K. Ohlendieck, Proteomics 9 (2009) 5509-5524.
 43. A. Lombardi, E. Silvestri, F. Cioffi, R. Senese, A. Lanni, F. Goglia, P. de Lange, M. Moreno, J. Proteomics 72 (2009) 708-721.
 44. J. Gannon, P. Doran, A. Kirwan, K. Ohlendieck, Eur. J. Cell Biol. 88 (2009) 685-700.
 45. C. Lewis, K. Ohlendieck, Anal. Biochem. 404 (2010) 197-203.
 46. L. Staunton, K. Ohlendieck, Protein Pept. Lett. 19 (2012) 252-263.
 47. D.W. Maughan, J.A. Henkin, J.O. Vigoreaux, Mol. Cell. Proteomics 4 (2005) 1541-1549.
 48. M. Toigo, S. Donohoe, G. Sperrazzo, B. Jarrold, F. Wang, R. Hinkle, E. Dolan, R.J. Isfort, R. Aebersold, Mol. Biosyst. 1 (2005) 229-241.
 49. X. Zuo, L. Echan, P. Hembach, H.Y. Tang, K.D. Speicher, D. Santoli, D.W. Speicher, Electrophoresis 22 (2001) 1603-1615.
 50. Y.H. Lee, H.T. Tan, M.C. Chung, Proteomics 10 (2010) 3935-3956.
 51. K. Ohlendieck, J.M. Ervasti, J.B. Snook, K.P. Campbell, J. Cell Biol. 112 (1991) 135-148.
 52. K. Ohlendieck, Biochim. Biophys. Acta. 1283 (1996) 215-222.
 53. S. Schiaffino, C. Reggiani, Physiol. Rev. 91 (2011) 1447-1531.
 54. D. Pette, R.S. Staron, Histochem. Cell Biol. 115 (2001) 359-372.
 55. P. Donoghue, P. Doran, P. Dowling, K. Ohlendieck, Biochim. Biophys. Acta. 1752 (2005) 166-176.
 56. P. Donoghue, P. Doran, K. Wynne, K. Pedersen, M.J. Dunn, K. Ohlendieck, Proteomics 7 (2007) 3417-3430.
 57. V. Zhukarev, J.M. Sanger, J.W. Sanger, Y.E. Goldman, H. Shuman, Cell Motil. Cytoskeleton 37 (1997) 363-377.
 58. Y. Bassaglia, J. Cebrian, S. Covan, M. Garcia, J. Foucrier, J. Exp. Cell Res. 302 (2005) 221-232.
 59. J. Gannon, L. Staunton, K. O'Connell, P. Doran, K. Ohlendieck, Int. J. Mol. Med. 22 (2008) 33-42.
 60. P. Doran, P. Donoghue, K. O'Connell, J. Gannon, K. Ohlendieck, Proteomics 9 (2009) 989-1003.
 61. K. Ohlendieck, Front. Physiol. 2 (2011) 105.
 62. A.M. Distler, J. Kerner, C.L. Hoppel, Proteomics 8 (2008) 4066-4082.
 63. C. Ruiz-Romero, F.J. Blanco, Mol. Biosyst. 5 (2009) 1130-1142.
 64. E. Silvestri, A. Lombardi, D. Glinni, R. Senese, F. Cioffi, A. Lanni, F. Goglia, M. Moreno, P. de Lange, J. Integ. OMICS 1 (2011) 17-27.
 65. S. Da Cruz, P.A. Parone, J.C. Martinou, Expert Rev. Proteomics 2 (2005) 541-551.
 66. M. Elstner, C. Andreoli, U. Ahting, I. Tetko, T. Klopstock, T. Meitinger, H. Prokisch, Mol. Biotechnol. 40 (2008) 306-315.
 67. X. Chen, J. Li, J. Hou, Z. Xie, F. Yang, Expert Rev. Proteomics

- 7 (2010) 333-345.
68. I.E. Scheffler, *Mitochondrion* 1 (2001) 3-31.
 69. H.M. McBride, M. Neuspiel, S. Wasiak, *Curr. Biol.* 16 (2006) R551-R560.
 70. M. Picard, T. Taivassalo, G. Gouspillou, R.T. Hepple, *J. Physiol.* 589 (2011) 4413-4421.
 71. J.M. Herrmann, S. Longen, D. Weckbecker, M. Depuydt, *Adv. Exp. Med. Biol.* 748 (2012) 41-64.
 72. O. Schmidt, N. Pfanner, C. Meisinger, *Nat. Rev. Mol. Cell Biol.* 11 (2010) 655-667.
 73. R. Ferreira, R. Vitorino, R.M. Alves, H.J. Appell, S.K. Powers, J.A. Duarte JA, F. Amado, *Proteomics* 10 (2010) 3142-3154.
 74. B.W. Kim, J.W. Lee, H.J. Choo, C.S. Lee, S.Y. Jung, J.S. Yi, Y.M. Ham, J.H. Lee, J. Hong, M.J. Kang, S.G. Chi, S.W. Hyung, S.W. Lee, H.M. Kim, B.R. Cho, D.S. Min, G. Yoon, Y.G. Ko, *Proteomics* 10 (2010) 2498-2515.
 75. H. Zhang, Y. Wang, J. Li, J. Yu, J. Pu, L. Li, H. Zhang, S. Zhang, G. Peng, F. Yang, P. Liu, *J. Proteome Res.* 10 (2011) 4757-4768.
 76. J. Feng, H. Xie, D.L. Meany, L.V. Thompson, E.A. Arriaga, T.J. Griffin, *J. Gerontol. A Biol. Sci. Med. Sci.* 63 (2008) 1137-1152.
 77. D.L. Meany, H. Xie, L.V. Thompson, E.A. Arriaga, T.J. Griffin, *Proteomics* 7 (2007) 1150-1163.
 78. D.C. Chan, *Cell* 125 (2006) 1241-1252.
 79. P.A. Figueiredo, M.P. Mota, H.J. Appell, J.A. Duarte, *Biogerontology* 9 (2008) 67-84.
 80. A. Safdar, M.J. Hamadeh, J.J. Kaczor, S. Raha, J. Debeer, M.A. Tarnopolsky, *PLoS One* 5 (2010) e10778.
 81. C.M. Peterson, D.L. Johannsen, E. Ravussin, *J. Aging Res.* 2012 (2012) 194821.
 82. L. Staunton, K. O'Connell, K. Ohlendieck, *J. Aging Res.* 2011 (2011) 908035.
 83. P.S. Brookes, A. Pinner, A. Ramachandran, L. Coward, S. Barnes, H. Kim, V.M. Darley-Usmar, *Proteomics* 2 (2002) 969-977.
 84. J. Gannon, K. Ohlendieck, *Mol. Med. Report* 5 (2012) 993-1000.
 85. A.A. Vandervoort, *Muscle Nerve* 25 (2002) 17-25.
 86. L.V. Thompson, *Exp. Gerontol.* 44 (2009) 106-111.
 87. D.R. Thomas, *Clin. Geriatr. Med.* 26 (2010) 331-346.
 88. M.J. Berger, T.J. Doherty, *Interdiscip. Top. Gerontol.* 37 (2010) 94-114.
 89. J.G. Burniston, *Biochim. Biophys. Acta.* 1784 (2008) 1077-1086.
 90. Y. Sato, M. Shimizu, W. Mizunoya, H. Wariishi, R. Tatsumi, V.L. Buchman, Y. Ikeuchi, *Biosci. Biotechnol. Biochem.* 73 (2009) 1748-1756.
 91. R. Ferreira, R. Vitorino, M.J. Neuparth, H.J. Appell, J.A. Duarte, F. Amado, *Eur. J. Appl. Physiol.* 107 (2009) 553-563.
 92. L. Staunton, H. Jockusch, K. Ohlendieck, *Biochem. Biophys. Res. Commun.* 406 (2011) 595-600.
 93. B. Egan, P. Dowling, P.L. O'Connor, M. Henry, P. Meleady, J.R. Zierath, D.J. O'Gorman, *Proteomics* (2011) 1413-1428.
 94. H. Andersen, S. Nielsen, C.E. Mogensen, J. Jakobsen, *J. Diabetes* 53 (2004) 1543-1548.
 95. H. Hwang, B.P. Bowen, N. Lefort, C.R. Flynn, E.A. De Filipis, C. Roberts, C.C. Smoke, C. Meyer, K. Højlund, Z. Yi, L.J. Mandarino, *Diabetes* 59 (2010) 33-42.
 96. J. Gieselstein, G. Poschmann, K. Højlund, W. Schechinger, J.W. Dietrich, K. Levin, H. Beck-Nielsen, K. Podwojski, K. Stuehler, H.E. Meyer, H.H. Klein, *Diabetologia* 55 (2012) 1114-11127.
 97. E. Mullen, K. Ohlendieck, *J. Integr. OMICS* 1 (2011) 108-114.
 98. K.P. Campbell, J.T. Stull, *J. Biol. Chem.* 278 (2003) 12599-12600.
 99. R.H. Aebersold, J. Leavitt, R.A. Saavedra, L.E. Hood, S.B. Kent, *Proc. Natl. Acad. Sci. USA* 84 (1987) 6970-6997.
 100. J.L. Luque-Garcia, G. Zhou, T.T. Sun, T.A. Neubert, *Anal. Chem.* 78 (2006) 5102-5108.
 101. J.L. Luque-Garcia, T.A. Neubert, *Methods Mol. Biol.* 536 (2009) 331-341.
 102. A.E. Rossi, R.T. Dirksen, *Muscle Nerve* 33 (2006) 715-731.
 103. D. Rossi, V. Barone, E. Giacomello, V. Cusimano, V. Sorrentino, *Traffic* 9 (2008) 1044-1049.
 104. B.E. Murray, G.R. Froemming, P.B. Maguire, K. Ohlendieck, *Int. J. Mol. Med.* 1 (1998) 677-687.
 105. C.L. Huang, T.H. Pedersen, J.A. Fraser, *J. Muscle Res. Cell. Motil.* 32 (2011) 171-202.
 106. V. Sorrentino, *Int. J. Biochem. Cell Biol.* 43 (2011) 1075-1078.
 107. E.M. Capes, R. Loaiza, H.H. Valdivia, *Skelet. Muscle* 1 (2011) 18.
 108. K.Y. Xu, L.C. Becker, *J. Histochem. Cytochem.* 46 (1998) 419-427.
 109. K.P. Campbell, D.H. MacLennan, A.O. Jorgensen, *J. Biol. Chem.* 258 (1983) 11267-11273.
 110. P. Doran, P. Dowling, J. Lohan, K. McDonnell, S. Poetsch, K. Ohlendieck, *Eur. J. Biochem.* 271 (2004) 3943-3952.
 111. T.A. Partridge, *Curr. Opin. Neurol.* 24 (2011) 415-422.
 112. F.W. Hopf, P.R. Turner, R.A. Steinhardt, *Subcell. Biochem.* 45 (2007) 429-464.
 113. P. Dowling, P. Doran, K. Ohlendieck, *Biochem J.* 379 (2004) 479-488.
 114. P. Doran, S.D. Wilton, S. Fletcher, K. Ohlendieck, *Proteomics* 9 (2009) 671-685.
 115. K. Ohlendieck, *Biochim. Biophys. Acta.* 1804 (2010) 2089-2101.
 116. Y. Ge, M.P. Molloy, J.S. Chamberlain, P.C. Andrews, *Proteomics* 3 (2003) 1895-1903.
 117. Y. Ge, M.P. Molloy, J.S. Chamberlain, P.C. Andrews, *Electrophoresis* 25 (2004) 2576-2585.
 118. P. Doran, P. Dowling, P. Donoghue, M. Buffini, K. Ohlendieck, *Biochim. Biophys. Acta.* 1764 (2006) 773-785.
 119. P. Doran, G. Martin, P. Dowling, H. Jockusch, K. Ohlendieck, *Proteomics* 6 (2006) 4610-4621.
 120. M. Stastna, J.E. Van Eyk, *Proteomics* 12 (2012) 722-735.
 121. C.Y. Chan, J.C. McDermott, K. W. Siu, *Int. J. Proteomics* 2011 (2011) 329467.
 122. X.C. Chan, J.C. McDermott, K.W. Siu, *J. Proteome Res.* 6 (2007) 698-710.
 123. J. Henningsen, K.T. Rigbolt, B. Blagoev, B.K. Pedersen, I. Kratchmarova, *Mol. Cell. Proteomics* 9 (2010) 2482-2496.
 124. F. Norheim, T. Raastad, B. Thiede, A.C. Rustan, C.A. Drevon, F. Haugen, *Am. J. Physiol. Endocrinol. Metab.* 301 (2011) E1013-E1021.
 125. J.H. Yoon, K. Yea, J. Kim, Y.S. Choi, S. Park, H. Lee, C.S. Lee, P.G. Suh, S.H. Ryu, *Proteomics* 9 (2009) 51-60.
 126. J.H. Yoon, P. Song, J.H. Jang, D.K. Kim, S. Choi, J. Kim, J. Ghim, D. Kim, S. Park, H. Lee, D. Kwak, K. Yea, D. Hwang, P.G. Suh, S.H. Ryu, *J. Proteome Res.* 10 (2011) 5315-5325.
 127. J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, *Nat. Methods* 6 (2009) 359-362.
 128. P.M. Donoghue, C. Hughes, J.P. Vissers, J.I. Langridge, M.J. Dunn, *Proteomics* 8 (2008) 3895-3905.
 129. P. Donoghue, L. Staunton, E. Mullen, G. Manning, K. Ohlendieck, *J. Proteomics* 73 (2010) 1441-1453.
 130. J.H. Yoon, E. Johnson, R. Xu, L.T. Martin, P.T. Martin, F. Montanaro, *J. Proteome Res.* 11 (2012) 4413-4424.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.77

Protein thiols as novel biomarkers in ecotoxicology: A case study of oxidative stress in *Mytilus edulis* sampled near a former industrial site in Cork Harbour, Ireland

Sara Tedesco¹, Siti Nur Tahirah Jaafar¹, Ana Varela Coelho² and David Sheehan^{1*}

¹Proteomics Research Group, Dept. Biochemistry and Environmental Research Institute, University College Cork, Western Gateway Building, Western Rd., Cork, Ireland. Telephone: +353 21 4205424. Fax number: +353 21 4205462. ²ITQB-UNL, Av. da República, Estação Agronómica Nacional, 2780-157 Oeiras, Portugal.

Received: 11 November 2011 Accepted: 01 May 2012 Available Online: 01 May 2012

ABSTRACT

Oxidative stress produces reactive oxygen species which can modify proteins and thiols of cysteines are especially susceptible. *Mytilus edulis* was sampled from three stations in Cork Harbour, Ireland and from an out-harbour control site in Bantry Bay, Ireland. A variety of traditional biomarkers were benchmarked against thiol oxidation. Lysosomal membrane stability diminished in haemocytes from the three Cork harbour sites, although a stronger effect was observed in two in-harbour stations of environmental concern (Douglas and Haulbowline Island). Catalase and glutathione transferase (GST) activities were decreased in digestive gland extracts of animals from in-harbour sites especially the in-harbour control (Ringaskiddy) showed lower GST than Bantry. Mussels collected at Haulbowline Island showed elevated lipid peroxidation ($p < 0.05$) compared to the other three stations and decreased levels of protein thiols which is consistent with oxidative stress at this site. Protein profiles for thiol-containing protein sub-proteomes trapped on activated thiol sepharose for each site were obtained by two dimensional electrophoresis and revealed differences between stations. Selected thiol-containing proteins were also identified by in-gel tryptic digestion and mass spectrometry; endoglucanase, aginine kinase, creatine kinase 1 and endo-1,4-beta-glucanase. Our findings confirmed that protein thiols are therefore sensitive novel biomarkers to oxidative stress.

Keywords: Protein thiols; biomarkers; remediation; oxidative stress; proteomics; *Mytilus edulis*.

1. Introduction

Cork Harbour, the second-busiest commercial port in the Republic of Ireland, is one of the world's largest natural harbours with a semi-enclosed area of approximately 25 km² [1] (Fig. 1). The harbour accepts anthropogenic inputs from industry, shipping, agricultural run-off and human sewage from the surrounding catchment of some 400,000 inhabitants and the underlying geology makes the water-body especially susceptible to pollution [2]. The intertidal area is an internationally-important wetland site for wintering waterfowl and is designated as a special protection area under the EU Birds Directive [3]. Although not extensively polluted by international standards [4], there is some localised build-up of PAHs, especially at Douglas estuary [5, 6, 7]. There is in-

tense local and national concern that industrial activities pose an ongoing threat to the quality of the aquatic environment of the harbour and a former steel plant on Haulbowline Island is the site of a major industrial remediation project. We have previously used protein biomarkers such as glutathione transferases (GSTs) [8, 9] and heat shock proteins [9] to assess the environmental stress-status of *Mytilus edulis* sampled from the harbour. Comet assays revealed PAH-mediated genotoxicity from sediment sampled around the harbour in turbot and clam [10]. More recently, we have used proteomic methods [11, 12] to extend these studies [13-16]. In the present investigation we have explored a novel redox proteomic method based on oxidation of protein thi-

*Corresponding author: Professor David Sheehan, Proteomics Research Group, Dept. Biochemistry and Environmental Research Institute, University College Cork, Western Gateway Building, Western Rd., Cork, Ireland. Telephone: +353 21 4205424. Fax number: +353 21 4205462. E-mail Address: d.sheehan@ucc.ie

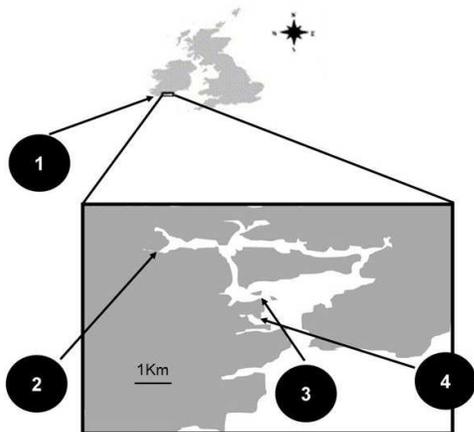


Figure 1. Sampling sites for *Mytilus edulis* around Cork harbour, Ireland: (1) Bantry Bay (out-harbour control site), (2) Douglas, (3) Haulbowline, (4) Ringaskiddy (in-harbour control site).

ols due to oxidative stress benchmarked against traditional biomarkers to assess the quality of mussels at sites adjacent to Haulbowline Island. We compared these to a reference site within the harbour (Ringaskiddy) and in Bantry Bay, a deep near-pristine inlet further along the West Cork coast. Our findings suggest that mussels from the two in-harbour test sites are under some environmental stress.

2. Materials and methods

2.1 Site selection and animals

M. edulis were collected from four sites in Ireland: Bantry Bay (9° 30'W, 51°40'N) and three sites within Cork Harbour; (Fig.1): Douglas (8° 23'W, 51°52'N), Haulbowline (8°17'W, 51°50'N), Ringaskiddy (8°18'W, 51°49'N). Bantry Bay is a clean area considered an appropriate control [17] which is an important site for commercial mussel aquaculture. The three Cork Harbour sites were chosen, respectively, because of presence of PAHs [Douglas; 5-7], the presence of large amounts of iron and other pollutants at a former steel plant presently undergoing remediation (Haulbowline) and previous history as an appropriate in-harbour control (Ringaskiddy; [9]). Thirty mussels (5-6 cm in length) were collected on a single day from the intertidal zone at low tide at the four sampling sites. Haemolymph was withdrawn using a 21 Gauge syringe from the adductor muscle of 10 animals for each site on the day of collection to measure lysosomal membrane stability in the haemocytes [18]. Digestive gland tissues were dissected, pooled (5 organisms per replicate), frozen in liquid nitrogen and stored at -80 °C.

2.2 Chemicals

Acetonitrile, bovine serum albumin (BSA), 1-chloro-2,4-dinitrochlorobenzene (CDNB), dimethyl sulfoxide (DMSO), 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB), 5'-

iodoacetamide fluorescein (IAF), 1-methyl-2-phenylindole, neutral red, phenyl-methylsulphonyl fluoride (PMSF), reduced glutathione (GSH), 1,1,3,3-tetramethoxypropane, and trifluoroacetic acid (TFA) were obtained from Sigma Chemical Co. (Poole, Dorset, UK). Activated thiol sepharose (ATS) was purchased from GE Healthcare (Little Chalfont, Bucks, UK).

2.3 Sample preparation

Digestive glands were homogenized in a motor-driven Teflon Potter-Elvehjem homogenizer in 10 mM Tris H-Cl, pH 7.2, containing 500 mM sucrose, 1 mM EDTA and 1 mM PMSF. Supernatants were collected by centrifugation at 20,000 ×g and stored at -70 °C until required for analysis. Protein content was calculated by the method of Bradford using BSA as a standard [19].

2.4 Lysosomal membrane stability

Lysosomal membrane stability was measured by the neutral red retention time assay [NRRT; 18]. Haemocytes from the adductor muscle were incubated on a glass slide with a freshly-prepared neutral red (NR) working solution (2 µl/ml saline from a stock solution of 20 mg neutral red dye dissolved in 1 ml of DMSO) and microscopically examined at 15 min intervals to determine the time at which 50% of cells had lost to the cytosol the dye previously taken up by lysosomes.

2.5 Antioxidant enzymes

Catalase activity (CAT) was measured according to the method of Aebi [20]. This method is based on measuring decrease in absorbance at 240 nm due to the consumption of hydrogen peroxide (H₂O₂). Activity was expressed as U/min/mg protein $\epsilon = -0.04 \text{ mM}^{-1} \text{ cm}^{-1}$). GST activity was determined using CDNB as substrate [21]. The reaction rate was detected at 340 nm, and expressed as nmol CDNB conjugate formed/ min/mg protein $\epsilon = 9.6 \text{ mM}^{-1} \text{ cm}^{-1}$.

2.6 Lipid peroxidation

Lipid peroxidation was measured by determining malondialdehyde (MDA), a decomposition product of polyunsaturated fatty acids. It was determined in samples homogenized (1:3 w/v) in 20 mM Tris-HCl pH 7.4, centrifuged at 3,000 ×g for 20 min and then derivatized in a 1 ml reaction mixture containing 10.3 mM 1-methyl-2-phenylindole (dissolved in acetonitrile/methanol 3:1), HCl 32%, 100 µl water and an equal volume of sample or standard (standard range 0–6 µM 1,1,3,3-tetramethoxypropane, in 20 mM Tris-HCl, pH 7.4). The tubes were vortexed and incubated at 45 °C for 40 min. Samples were cooled on ice, centrifuged at 15,000 ×g for 10 min and read spectrophotometrically at 586 nm; levels of MDA were calibrated against an MDA

standard curve and expressed as nmol/g wet weight [22].

2.7 Labelling protein thiols

Protein thiols present in protein extracts were labelled by adding IAF in DMSO to a final concentration of 800 μ M and incubating at room temperature for 2 h in the dark. IAF reacts specifically with reduced thiols (-SH) but not with oxidised variants such as sulphenic acid (-SOH) or disulphides (-S-S-) which might be expected to form on oxidative stress [23].

2.8 Protein electrophoresis

Proteins were resolved using one-dimensional electrophoresis (1DE) in 12% polyacrylamide gels [24]. Samples were diluted in buffer lacking β -mercaptoethanol, to avoid reduction of disulphide bridges. Gels were scanned in a Typhoon 9400 scanner (GE Healthcare, UK; excitation, 490-495 nm; emission, 515-520 nm) and were subsequently stained with Coomassie G250. Equal amounts of protein were loaded in 12 wells (3 replicates for each treatment) and repeated at least 3 times. Two-dimensional SDS PAGE electrophoresis (2DE) analysis was performed on protein extracts trapped by covalent disulphide exchange on ATS [25-27]. Protein samples were precipitated with TCA/acetone and re-suspended in rehydration buffer containing 5M urea, 2 M thiourea, 2% CHAPS, 4% ampholyte (Pharmalyte 3-10, Amersham-Pharmacia Biotech, Little Chalfont, Bucks., UK), 1% Destreak reagent (Amersham-Pharmacia Biotech) and trace amounts of bromophenol blue. A final volume of 125 μ l was loaded on 7cm IPG strips pH 3-10NL on the bench overnight. Proteins were focused on a Protean IEF Cell (Biorad) with linear voltage increases: 250 V for 15 min: 4000 V for 2 h; then up to 20,000 Vh. After focusing, strips were equilibrated for 15 min in equilibration buffer (6 M urea, 0.375 M Tris, pH 8.8, 2% SDS, 20% glycerol) containing 2% DTT and then for 15 min in equilibration buffer containing 2.5% iodoacetamide. Equilibrated strips were electrophoresed on 12% SDS-PAGE gels at a constant voltage (150 V) at 4 $^{\circ}$ C. 2D gels were scanned by calibrated densitometer (Bio-Rad Laboratories) of gels visualized by Colloidal Coomassie staining [28].

2.9 Quantification of proteins

For each 1DE gel, bands detected by the Typhoon 9400 scanner, were subsequently analyzed by Quantity One image analysis software (Bio-Rad, Hercules, CA, USA) measuring the total intensity for each lane, quantified as arbitrary units (A.U.). 1DE gels stained with Coomassie blue G250 were scanned in a GS-800 calibrated densitometer and total optical density of each lane measured by Quantity One image analysis software. Total optical densities for each lane were normalized with those from coomassie staining for the same gel track.

2.10 In-gel digestion and MALDI-TOF/TOF analysis

Proteins were excised from gels and cleaved with trypsin by in-gel digestion. The protein spot digestion was performed in OMX-S devices according to the manufacturer's instructions (OMX, Munich, Germany). Briefly, 20 μ l of modified trypsin (10 ng/ μ L in 50mM ammonium bicarbonate) were added to the device and centrifuged briefly at 3,800 x g. The digestion procedure was carried out at 50 $^{\circ}$ C for 45 minutes with gentle agitation. The peptide solution was removed from the reactor compartment by centrifugation at 1,000 x g for 3 min. Peptide solutions were desalted and concentrated with chromatographic microcolumns using GELoader tips packed with POROS R2 (Applied Biosystems, Foster City, California, USA; 20 μ m bead size) and then directly eluted onto the MALDI target plate using 0.5 μ l of 5 mg/ml α -ciano-4-hydroxy-trans-cinnamic acid (α -CHCA) in 50% (v/v) ACN with 2.5% (v/v) formic acid and air-dried.

Tandem mass spectrometry analysis was performed using a MALDI-TOF/TOF 4800plus mass spectrometer (Applied Biosystems). The equipment was calibrated using angiotensin II (1,046.542 Da), angiotensin I (1,296.685 Da), Neurotensin (1,672.918 Da), adrenocorticotrophic hormone (ACTH) (1-17) (2,093.087 Da), and ACTH (18-39) (2,465.199) (Peptide Calibration Mixture 1, LaserBio Labs, Sophia-Antipolis, France). Each reflector MS spectrum was collected in a result-independent acquisition mode, typically using 750 laser shots per spectra and a fixed laser intensity of 3,200V. The fifteen strongest precursors were selected for MS/MS, the strongest precursors being fragmented first. MS/MS analyses were performed using collision induced dissociation (CID) assisted with air, with collision energy and gas pressure of 1 kV and 1 x 10⁶ torr, respectively. Each MS/MS spectrum collected consisted of 1,200 laser shots using a fixed laser intensity of 4,300V.

2.11 Protein identification

Protein identification was performed using MASCOT (version 2.2; Matrix Science, Boston, MA) search engine. Searches were performed using combined analysis of the intact masses of the tryptic peptides (MS) and tandem mass data (MS/MS). Search parameters were set as follows: minimum mass accuracy of 50 ppm for the parent ions, an error of 0.3 Da for the fragments, two missed cleavages in peptide masses, and carbamidomethylation (C), oxidation (M), deamidation (NQ), Gln->pyro-Glu (N-term Q) were set as variable amino acid modifications and a non-redundant NCBI database (released 2012_01) was used. Peptides were only considered if the ion score indicated extensive homology (p<0.05). Proteins were considered if having significant MASCOT score and at least one peptide with extensive sequence homology. Automated GO annotation was performed using the GO categories of the best hit derived from the BLASTp results (BLASTp minimal expectation value set

to 1×10^{-3}) for additional information on functional pathway.

2.12 Statistical analysis

All data are means \pm standard deviation (SD) of triplicate determinations on three independent extracts for each treatment studied. Statistical analyses of data were performed using the Software Statistica 7.0 (Stat Soft, Tulsa, Oklahoma, USA). Samples were tested using one-way ANOVA, homogeneity of variance was tested by Cochran C and mathematical transformation applied if necessary; post hoc comparison (Newman–Keuls) was used to discriminate between means of values. Differences were considered statistically significant when $p < 0.05$.

3. Results

3.1 Lysosomal membrane stability

Lysosomal membrane stability measured in freshly-sampled haemocytes showed lower NRRT at all three stations compared to the out-harbour control site, Bantry Bay. The NRRT observed in mussels from Douglas and Haulbowline were less than half that for Bantry while Ringaskiddy was approximately 80% of Bantry (Fig.2).

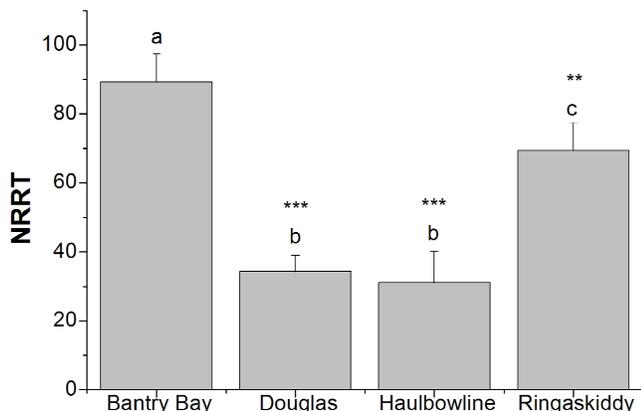


Figure 2. Lysosomal membrane stability measured as neutral red retention time (NRRT; min) in haemolymph from *M. edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Data are expressed as mean \pm SD. Superscripts of different letters are significantly different from each other at ** $p < 0.01$; *** $p < 0.005$.

3.2 Antioxidant enzymes

CAT activity was significantly lower ($p < 0.05$) in mussels collected from the three in-harbour stations compared to Bantry Bay (Fig.3).

Significantly lower GST activity was found in mussel digestive glands at the three in-harbour stations compared to Bantry Bay (Fig.4).

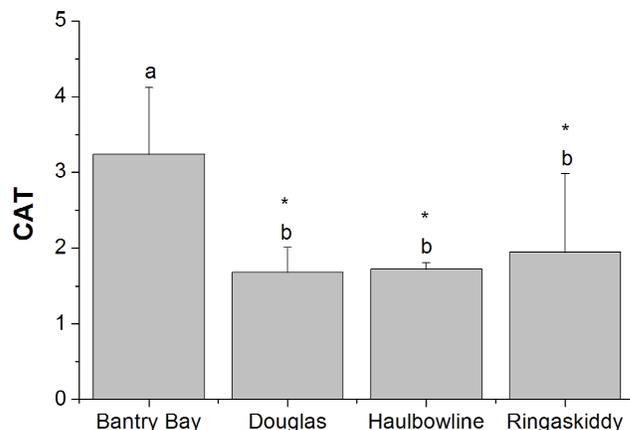


Figure 3. Catalase activities (CAT; U/min/mg prot) in digestive gland of *M. edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Data are expressed as mean \pm SD. Superscripts of different letters are significantly different from each other at * $p < 0.05$.

3.3 Lipid peroxidation

The highest MDA levels were determined in digestive gland of mussels collected at Haulbowline, showing higher lipid peroxidation ($p < 0.05$) compared to the out-harbour control site, Bantry Bay (Fig.5).

3.4 Protein thiols

1DE separation of IAF-labelled proteins revealed a decrease in total thiol-containing proteins in samples from

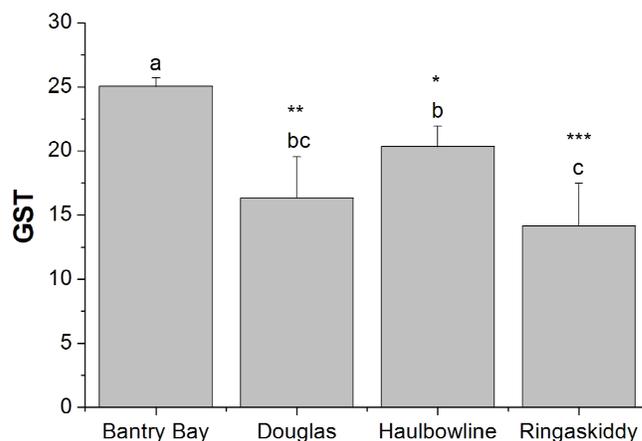


Figure 4. Glutathione transferase activities (GST; nmol CDNB conjugate formed/min/mg prot) in digestive gland of *M. edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Data are expressed as mean \pm SD. Superscripts of different letters are significantly different from each other at * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$.

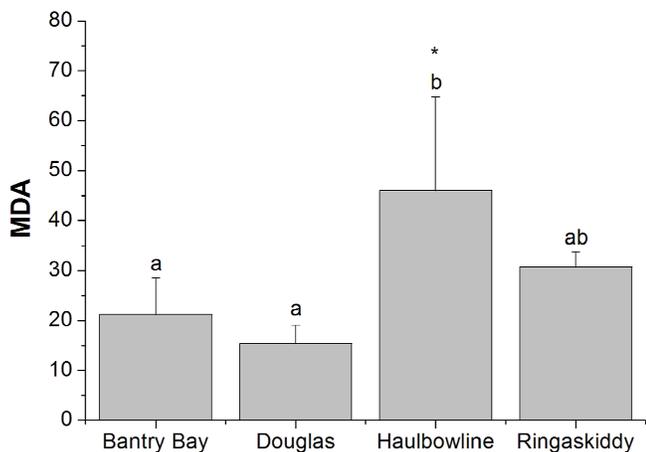


Figure 5. Lipid peroxidation measured as malondialdehyde (MDA) levels (nmol/g wet weight tissue) in digestive gland of *M.edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Data are expressed as mean ± SD. Superscripts of different letters are significantly different from each other at *p<0.05.

Haulbowline (p=0.03) and Douglas (p=0.052) compared to Bantry Bay (Fig.6).

Thiol-containing proteins were trapped on ATS [25-27] and separated by 2DE (Fig. 7). This revealed differences in spot patterns which we attribute to oxidation of thiols in specific proteins supporting the results obtained by 1DE. Taken together, these data suggest that protein thiols decreased strongly in digestive gland of mussels collected in Haulbowline followed by Douglas.

3.5 Identified proteins

The weakened or missing spots from 2DE of samples col-

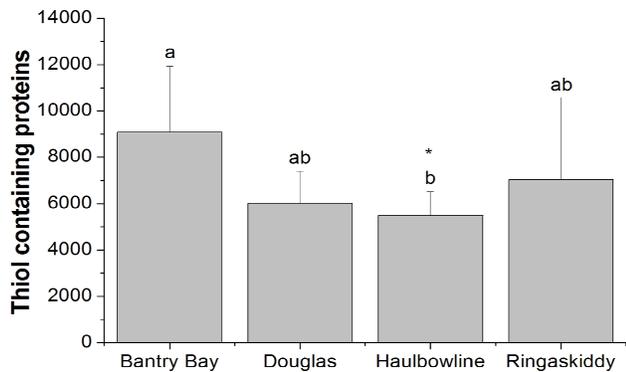


Figure 6. 1DE of thiol-containing proteins (normalized by total protein amounts) in digestive gland of *M. edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Data are expressed in A.U. as mean ± SD. Superscripts of different letters are significantly different from each other at *p<0.05.

lected in Haulbowline and Douglas were selected for protein identification (Fig.7). The reason is that these thiol-containing proteins were unable to be selected by ATS due to oxidation of cysteines. Because of the lack of a full genome sequence for *Mytilus* species it was necessary to search other species for matches with peptides derived from in-gel digestion of selected spots. Four of the selected proteins were successfully identified in this way showing significant MASCOT scores and at least one peptide with extensive sequence homology (Table 1).

The MS results confirmed that all selected spots were thiol-containing proteins but only one of them, endo-1,4-beta-glucanase (spot 4), was found to correspond to an *M. edulis* sequence. Spot 1 was identified as another endoglucanase but from *Mizuhopecten yessoensis*, a scallop belonging to the family of Pectinidae but to the same class of Bivalvia as *M. edulis*. Spot 2 matched well with arginine kinase from *Conus novaehollandiae*, a marine gastropod mollusc belonging to the same Phylum (Mollusca) as *M. edulis*. Spot 3 was similar to creatine Kinase 1 of *Lethenteron camtschaticum*, a freshwater fish not taxonomically similar to the blue mussel but included because of its high MASCOT score and good expectation value (Table 1).

4. Discussion

Cysteine is the second least-abundant residue in proteins and is the main point of crosstalk between redox status and cell signalling [29]. Both in controlled exposure experiments with pro-oxidants in holding tanks [30] and in the field [31], we have previously shown that mussels change aspects of their thiol chemistry in response to oxidative stress. Because of their roles in buffering transient increases in ROS [32] and in cellular redox signaling pathways [29], protein thiols are especially attractive targets as possible novel biomarkers for oxidative stress. A multiplexing approach allowed simultaneous determination of total thiols and total protein in electrophoretic separations by exploiting the specificity of IAF for reduced thiols. This was performed in this study and decreased IAF labeling is attributed to thiol oxidation [23]. The Haulbowline site showed decreased total thiols suggesting that proteins, as well as lipids, experienced attack by ROS at this site. This observation was extended by trapping thiol-containing proteins on ATS and analyzing the thiol-proteome by 2DE [26-28]. This revealed closely-comparable separations in samples from each site but with individual spot differences, consistent with differences in thiol status across the sub-proteomes. *M. edulis* is a sentinel species widely-used in surveillance of environmental quality with particular relevance to marine estuaries [33, 34]. A number of biomarkers useful for assessing environmental quality have been developed in mussels [35, 36] but we are interested in identifying novel protein biomarkers that may complement these traditional indices and possibly yield greater insights to toxicity mechanisms [9, 13-17]. In the present case study, two sites of environmental interest within Cork Har-

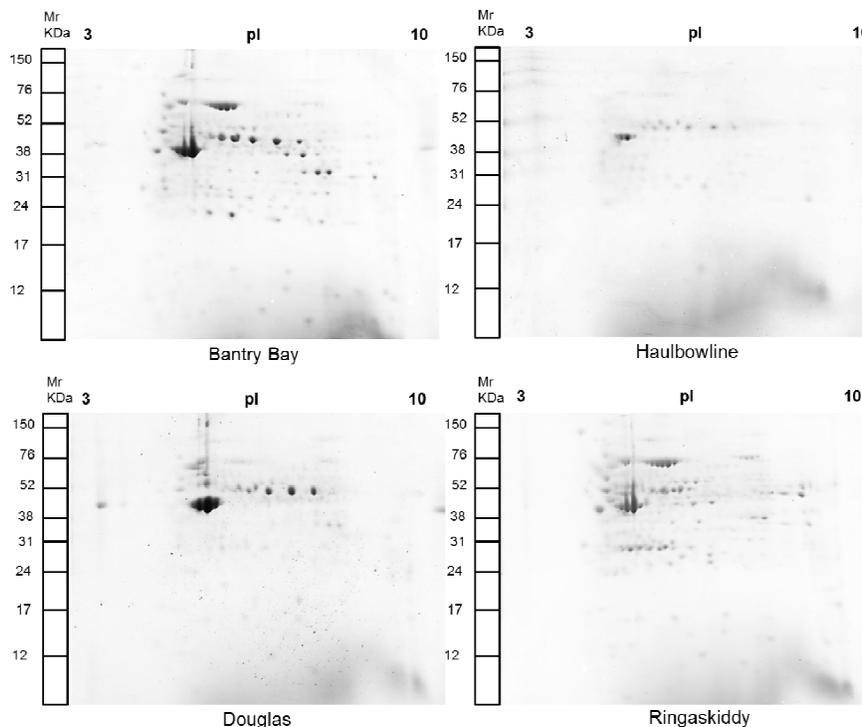


Figure 7. Representative 2DE of thiol containing proteins trapped by ATS in digestive gland of *M. edulis* collected in Bantry Bay (out-harbour control site), Douglas and Haulbowline (polluted sites) and Ringaskiddy (in-harbour control site). Spots present in Bantry Bay and/or Ringaskiddy 2DE but weakened or missing in Douglas and Haulbowline are shown with arrows. These spots were selected for the protein identification.

Table 1. Identified proteins in digestive gland of *M. edulis* by MALDI TOF MS. Cut-off MASCOT score ($p < 0.05$) = 84.

Spot #	Accession #	Protein ID	Organism	Theoretical/observed pI	Theoretical/observed Mr (kDa)	Mascot score	Expectation value	Peptides identified	Function
1	gi 254553092	Endoglucanase	<i>Mizuhopecten yessoensis</i>	5,66/~5.5	64.152/52-76	120	1.5e-0005	1	Carbohydrate metabolism
2	gi 301341836	Arginine kinase	<i>Conus novae-hollandiae</i>	6,34/~6.4	39.379/31-38	135	4.8e-007	1	Energy metabolism
3	gi 42627683	Creatine kinase 1	<i>Lethenteron camtschaticum</i>	6,71/~6.8	42.620/31-38	166	3.8e-010	1	Energy metabolism
4	gi 12230122	Endo-1,4-beta-glucanase	<i>Mytilus edulis</i>	6,79/~6	19.699/17-24	582	9.6e-052	5	Carbohydrate metabolism

bour, Ireland, were compared to in-harbour and out-harbour reference sites by measuring physiological and biochemical indices of stress, mainly in digestive gland, a key site of enzymatic detoxification [37]. Haemocyte NRRT was decreased in the two in-harbour test sites while lipid peroxidation was significantly increased in the Haulbowline site. This is consistent with significant environmental stress at that location possibly arising from residual pollutants from recently-removed slag heaps in a former steel plant currently undergoing remediation.

Lower levels of CAT and GST activities at both sites are consistent with these findings since these enzymes contribute strongly to defence against xenobiotic and oxidative

stress [38-40]. CAT detoxifies H_2O_2 to water *in vivo* but, in the presence of iron, H_2O_2 produces the hydroxyl radical by means of the Fenton reaction [29]. H_2O_2 is formed naturally in water as a result of photo-oxidation [41]. Enhanced CAT activity has been reported in fish and invertebrate species [42, 43] and its inhibition has been suggested as a transitory response to acute pollution [36]. Decreased CAT activity, combined with available iron, could potentially result in formation of the hydroxyl radical leading to oxidative stress.

GST is a phase II detoxification enzyme that catalyses conjugation of electrophiles to GSH [40] which has also found use as a biomarker in mussels [9, 17]. In this study, all three Cork Harbour sites showed slightly lower GST levels

than Bantry Bay. GST contributes to protecting tissues from oxidative stress by catalytic detoxification and binding as it is inducible by a wide range of chemical agents, some of which are also substrates such as hydroperoxides [40]. Increase of GST activity can therefore be due to increased detoxification of hydroperoxides. GSH conjugates are subsequently enzymatically degraded to mercapturates and excreted. However, under more intense or prolonged oxidative stress conditions, GST catalytic activity can be compromised due to conjugation of GSH to xenobiotic electrophilic centres causing a depletion of GSH and GST inhibition [44]. Generally, GST activity is lower in mussel digestive gland than gill [17]. We have previously found that gill GST activity varies with pollution status [9, 17]. However, while GST activity of *M. edulis* digestive gland did vary between sites on the South Coast of Ireland [17] it did not vary in digestive gland of *M. galloprovincialis* sampled from polluted sites in Venice Lagoon [8]. This suggests that digestive gland GST level may not be as responsive to pollution as gill.

MDA level is one of the oxidative stress parameters that has been measured in bivalve molluscs, especially in mussels, to investigate the biomarker's response to cellular free radical toxicity under metal exposure [45]. Digestive glands of mussels collected at Haulbowline showed statistically higher MDA levels than those sampled from Bantry Bay ($p=0.046$) but not versus Ringaskiddy (in-harbour control). Intriguingly, no relevant effects of lipid peroxidation or decrease of amount of thiol-containing protein were observed in mussels collected in Douglas. Our results showed that MDA levels in digestive gland of mussels collected at Haulbowline were statistically higher than those sampled from Bantry Bay ($p=0.046$) but not versus Ringaskiddy (in-harbour control). Intriguingly, no relevant effects of lipid peroxidation or decrease of amount of thiol-containing proteins was observed in mussels collected in Douglas.

These data suggest that protein thiols play a role in protection against lipid peroxidation. Oxidation of protein thiols, usually occurs by two different mechanisms: (1) lipid peroxidation induced by the depletion of GSH generates reactive aldehydes [46, 47] which may react with protein thiols; (2) reactive metabolites may react directly with protein thiols. Metals and their chelate complexes, such as copper, chromium, nickel, and cadmium, are implicated in lipid peroxidation [48]. It is likely that organic compounds (e.g. PAH, PCB) or metals released to seawater from the former steel plant on Haulbowline Island affected the local intertidal *M. edulis* population. To our knowledge, this report is the first to show redox cysteine modifications in endoglucanase, arginine kinase and creatine kinase proteins in *M. edulis*.

Endoglucanases are enzymes belonging to the cellulase family involved in carbohydrate metabolism. The endoglucanases identified are rich in cysteine residues and endo-1,4-glucanase which has been sequenced and cloned in digestive gland of *M. edulis* [49,50] contains twelve cysteine residues involved in six disulfide bonds. It is thought that these disulphides may contribute to anti-freezing properties of this

protein at low temperature. Decrease of cellulase activity has been considered as one of the major potential biomarkers for exposure to pesticides in aquatic invertebrates [51]. However De Coen et al. (2001) have found that relationships between enzymatic endpoints in carbohydrate metabolism and population level effects observed in *Daphnia magna* were toxicant-specific, and no single enzyme in carbohydrate metabolism could predict quantitative changes in population characteristics [52].

Arginine kinase catalyzes the reversible transfer from phospho-L-arginine to ADP to form ATP and is important for buffering ATP levels during burst muscle contraction in invertebrates [53]. Our results showed that this enzyme was down-regulated in Douglas and especially Haulbowline. This enzyme has been proposed to increase the ability of invertebrates to cope with the stress of variable environmental conditions related to hypoxia and acidosis [54, 55]. Silvestre et al. (2006) found strong down-regulation of this enzyme in the gills of the Chinese mitten crab after chronic cadmium exposure [56]. We also found that creatine kinase is another thiol-containing protein that may be sensitive to oxidative stress. It is a key enzyme in energy metabolism catalyzing reversible phosphorylation of creatine by ATP [55]. Sethuraman et al. (2004) showed that cysteine thiols of sarcomeric creatine kinase were oxidized after exposure to high concentrations of hydrogen peroxide [57]. A previous study showed that arginine and creatine kinase are sensitive to oxidation although by different mechanisms. Mammalian creatine kinase was very sensitive to the superoxide radical resulting in loss of enzyme activity whereas arginine kinase was less affected by comparable exposure [58]. These authors also found that loss of creatine kinase activity can be due to its high susceptibility to hypoxic conditions. The increase of nutrients usually found in polluted areas like some harbours [12, 59, 60] can be the result of algal bloom and subsequent hypoxia and ROS production enhance.

Taken together, our results show that mussels from the Haulbowline site experience considerable physiological and oxidative stress, which is consistent with the presence of pollutants originating from the nearby former steel plant. We suggest that protein thiols may be a potentially useful novel biomarker for oxidative stress in environmental toxicology.

Acknowledgements

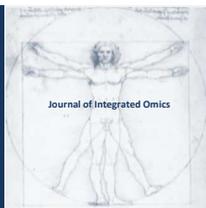
Our laboratory is supported by the Programme for Research in Third Level Institutions of the Higher Education Authority of Ireland. ST received an EMBARK award from the Irish Research Council for Science Engineering and Technology.

References

- [1] O'Mahony, C., Gault, J., Cummins, V., Köpke, K., O'Suilleabhain, D., Assessment of recreation activity and its application

- to integrated management and spatial planning for Cork Harbour, Ireland. *Marine Policy* 2009, 33, 930-937.
- [2] Allen, A., Milenic, D., Groundwater vulnerability assessment of the Cork Harbour area, SW Ireland. *Environmental Geology* 2007, 53, 485-492.
 - [3] EU, (1979) http://europa.eu/legislation_summaries/other/l28046_en.htmSite.
 - [4] Johnson, M. P., Hartnett, M., Collier, L. M., Costello, M. J., Identifying 'hot spots' of biological and anthropogenic activity in two Irish estuaries using means and frequencies. *Hydrobiologia* 2002, 475-476, 111-123.
 - [5] Byrne, P. A., Halloran, J. O., Aspects of Assaying Sediment Toxicity in Irish Estuarine Ecosystems. *Marine Pollution Bulletin* 1999, 39, 97-105.
 - [6] Kilemade, M., Hartl, M. G. J., Sheehan, D., Mothersill, C., et al., An assessment of the pollutant status of surficial sediment in Cork Harbour in the South East of Ireland with particular reference to polycyclic aromatic hydrocarbons. *Marine Pollution Bulletin* 2004, 49, 1084-1096.
 - [7] Kilemade, M., Hartl, M. G., O'Halloran, J., O'Brien, N. M., et al., Effects of contaminated sediment from Cork Harbour, Ireland on the cytochrome P450 system of turbot. *Ecotoxicology and environmental safety* 2009, 72, 747-755.
 - [8] Fitzpatrick, P. J., Sheehan, D., Livingstone, D. R., Studies on isoenzymes of glutathione S-transferase in the digestive gland of *Mytilus galloprovincialis* with exposure to pollution. *Marine Environmental Research* 1995, 39, 241-244.
 - [9] Lyons, C., Dowling, V., Tedengren, M., Gardestrom, J., et al., Variability of heat shock proteins and glutathione S-transferase in gill and digestive gland of blue mussel, *Mytilus edulis*. *Mar Environ Res* 2003, 56, 585-597.
 - [10] Hartl, M. G., Kilemade, M., Coughlan, B. M., O'Halloran, J., et al., A two-species biomarker model for the assessment of sediment toxicity in the marine and estuarine environment using the comet assay. *Journal of environmental science and health. Part A, Toxic/hazardous substances & environmental engineering* 2006, 41, 939-953.
 - [11] Sheehan, D., Detection of redox-based modification in two-dimensional electrophoresis proteomic separations. *Biochemical and biophysical research communications* 2006, 349, 455-462.
 - [12] Sheehan, D., McDonagh, B., Oxidative stress and bivalves: a proteomic approach. *Invertebrate Survival Journal* 2008, 5, 110-123.
 - [13] Dowling, V. A., Sheehan, D., Proteomics as a route to identification of toxicity targets in environmental toxicology. *Proteomics* 2006, 6, 5597-5604.
 - [14] McDonagh, B., Sheehan, D., Redox proteomics in the blue mussel *Mytilus edulis*: carbonylation is not a pre-requisite for ubiquitination in acute free radical-mediated oxidative stress. *Aquat Toxicol* 2006, 79, 325-333.
 - [15] McDonagh, B., Sheehan, D., Effects of oxidative stress on protein thiols and disulphides in *Mytilus edulis* revealed by proteomics: actin and protein disulphide isomerase are redox targets. *Mar Environ Res* 2008, 66, 193-195.
 - [16] McDonagh, B., Tyther, R., Sheehan, D., Redox proteomics in the mussel, *Mytilus edulis*. *Mar Environ Res* 2006, 62 Suppl, S101-104.
 - [17] Patrick J. Fitzpatrick, J. O. H. D. S., Andrew, R. W., Assessment of a glutathione S-transferase and related proteins in the gill and digestive gland of *Mytilus edulis* (L.), as potential organic pollution biomarkers. *Biomarkers* 1997, 2, 51-56.
 - [18] DM, L., VU, F., MH, D., Contaminant-induced lysosomal membrane damage in blood cells of mussels *Mytilus galloprovincialis* from the Venice Lagoon: an in vitro study. *Marine Ecology Progress Series* 1995, 129, 189-196.
 - [19] Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* 1976, 72, 248-254.
 - [20] Aebi, H., Methods of enzymatic analysis. *Catalase Estimation* 1974, 673-684.
 - [21] Habig, W. H., Pabst, M. J., Jakoby, W. B., Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *The Journal of biological chemistry* 1974, 249, 7130-7139.
 - [22] Shaw, J. P., Large, A. T., Donkin, P., Evans, S. V., et al., Seasonal variation in cytochrome P450 immunopositive protein levels, lipid peroxidation and genetic toxicity in digestive gland of the mussel *Mytilus edulis*. *Aquat Toxicol* 2004, 67, 325-336.
 - [23] Baty, J. W., Hampton, M. B., Winterbourn, C. C., Detection of oxidant sensitive thiol proteins by fluorescence labeling and two-dimensional electrophoresis. *Proteomics* 2002, 2, 1261-1266.
 - [24] Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970, 227, 680-685.
 - [25] Caldas, T. D., El Yaagoubi, A., Kohiyama, M., Richarme, G., Purification of elongation factors EF-Tu and EF-G from *Escherichia coli* by covalent chromatography on thiol-sepharose. *Protein expression and purification* 1998, 14, 65-70.
 - [26] Hu, W., Tedesco, S., McDonagh, B., Barcena, J. A., et al., Selection of thiol- and disulfide-containing proteins of *Escherichia coli* on activated thiol-Sepharose. *Analytical biochemistry* 2010, 398, 245-253.
 - [27] Hu, W., Tedesco, S., Faedda, R., Petrone, G., et al., Covalent selection of the thiol proteome on activated thiol sepharose: a robust tool for redox proteomics. *Talanta* 2010, 80, 1569-1575.
 - [28] Dyballa, N., Metzger, S., Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *Journal of visualized experiments : JoVE* 2009.
 - [29] Winterbourn, C. C., Reconciling the chemistry and biology of reactive oxygen species. *Nature chemical biology* 2008, 4, 278-286.
 - [30] McDonagh, B., Sheehan, D., Effect of oxidative stress on protein thiols in the blue mussel *Mytilus edulis*: proteomic identification of target proteins. *Proteomics* 2007, 7, 3395-3403.
 - [31] Prevodnik, A., Gardestrom, J., Lilja, K., Elfving, T., et al., Oxidative stress in response to xenobiotics in the blue mussel *Mytilus edulis* L.: evidence for variation along a natural salinity gradient of the Baltic Sea. *Aquat Toxicol* 2007, 82, 63-71.
 - [32] Hansen, R. E., Roth, D., Winther, J. R., Quantifying the global cellular thiol-disulfide status. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106, 422-427.
 - [33] Bayne, B. L., Mussel watching. *Nature* 1978, 275, 87-88.
 - [34] Goldberg, E. D., Bowen, V. T., Farrington, J. W., Harvey, G., et al., The Mussel Watch. *Environmental Conservation* 1978, 5, 101-125.
 - [35] Cajaraville, M. P., Bebianno, M. J., Blasco, J., Porte, C., et al., The use of biomarkers to assess the impact of pollution in coastal environments of the Iberian Peninsula: a practical approach. *The Science of the total environment* 2000, 247, 295-311.
 - [36] Regoli, F., Principato, G., Glutathione, glutathione-dependent and antioxidant enzymes in mussel, *Mytilus galloprovincialis*, exposed to metals under field and laboratory conditions: implications for the use of biochemical biomarkers. *Aquatic Toxicology* 1995, 31, 143-164.
 - [37] Power, A., Sheehan, D., Seasonal Variation in the Antioxidant

- Defence Systems of Gill and Digestive Gland of the Blue Mussel, *Mytilus edulis*. *Comparative Biochemistry and Physiology -- Part C: Pharmacology, Toxicology and Endocrinology* 1996, 114, 99-103.
- [38] Bocchetti, R., Fattorini, D., Pisanelli, B., Macchia, S., et al., Contaminant accumulation and biomarker responses in caged mussels, *Mytilus galloprovincialis*, to evaluate bioavailability and toxicological effects of remobilized chemicals during dredging and disposal operations in harbour areas. *Aquat Toxicol* 2008, 89, 257-266.
- [39] Frenzilli, G., Bocchetti, R., Pagliarecci, M., Nigro, M., et al., Time-course evaluation of ROS-mediated toxicity in mussels, *Mytilus galloprovincialis*, during a field translocation experiment. *Mar Environ Res* 2004, 58, 609-613.
- [40] Sheehan, D., Meade, G., Foley, V. M., Dowd, C. A., Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *The Biochemical journal* 2001, 360, 1-16.
- [41] Draper, W. M., Crosby, D. G., The photochemical generation of hydrogen peroxide in natural waters. *Archives of Environmental Contamination and Toxicology* 1983, 12, 121-126.
- [42] Di Giulio, R. T., Habig, C., Gallagher, E. P., Effects of Black Rock Harbor sediments on indices of biotransformation, oxidative stress, and DNA integrity in channel catfish. *Aquatic Toxicology* 1993, 26, 1-22.
- [43] Stephenson, J., Exposure to home pesticides linked to Parkinson disease. *JAMA : the journal of the American Medical Association* 2000, 283, 3055-3056.
- [44] Regoli, F., Trace metals and antioxidant enzymes in gills and digestive gland of the Mediterranean mussel *Mytilus galloprovincialis*. *Arch Environ Contam Toxicol* 1998, 34, 48-63.
- [45] Valavanidis, A., Vlahogianni, T., Dassenakis, M., Scoullou, M., Molecular biomarkers of oxidative stress in aquatic organisms in relation to toxic environmental pollutants. *Ecotoxicology and environmental safety* 2006, 64, 178-189.
- [46] Esterbauer, H., Benedetti, A., Lang, J., Fulceri, R., et al., Studies on the mechanism of formation of 4-hydroxynonenal during microsomal lipid peroxidation. *Biochimica et biophysica acta* 1986, 876, 154-166.
- [47] Pompella, A., Romani, A., Fulceri, R., Benedetti, A., Comporti, M., 4-Hydroxynonenal and other lipid peroxidation products are formed in mouse liver following intoxication with allyl alcohol. *Biochimica et biophysica acta* 1988, 961, 293-298.
- [48] Sole, J., Huguet, J., Arola, L., Romeu, A., In vivo effects of nickel and cadmium in rats on lipid peroxidation and ceruloplasmin activity. *Bulletin of environmental contamination and toxicology* 1990, 44, 686-691.
- [49] Xu, B., Hellman, U., Ersson, B., Janson, J. C., Purification, characterization and amino-acid sequence analysis of a thermostable, low molecular mass endo-beta-1,4-glucanase from blue mussel, *Mytilus edulis*. *European journal of biochemistry / FEBS* 2000, 267, 4970-4977.
- [50] Xu, B., Janson, J. C., Sellos, D., Cloning and sequencing of a molluscan endo-beta-1,4-glucanase gene from the blue mussel, *Mytilus edulis*. *European journal of biochemistry / FEBS* 2001, 268, 3718-3727.
- [51] Hyne, R. V., Maher, W. A., Invertebrate biomarkers: links to toxicosis that predict population decline. *Ecotoxicology and environmental safety* 2003, 54, 366-374.
- [52] De Coen, W. M., Janssen, C. R., Segner, H., The use of biomarkers in *Daphnia magna* toxicity testing V. In vivo alterations in the carbohydrate metabolism of *Daphnia magna* exposed to sublethal concentrations of mercury and lindane. *Ecotoxicology and environmental safety* 2001, 48, 223-234.
- [53] Livingstone, D. R., de Zwaan, A., Thompson, R. J., Aerobic metabolism, octopine production and phosphoarginine as sources of energy in the phasic and catch adductor muscles of the giant scallop *Placopecten magellanicus* during swimming and the subsequent recovery period. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 1981, 70, 35-44.
- [54] Ellington, W. R., Phosphocreatine represents a thermodynamic and functional improvement over other muscle phosphagens. *The Journal of experimental biology* 1989, 143, 177-194.
- [55] Ellington, W. R., EVOLUTION AND PHYSIOLOGICAL ROLES OF PHOSPHAGEN SYSTEMS. *Annual Review of Physiology* 2001, 63, 289-325.
- [56] Silvestre, F., Dierick, J. F., Dumont, V., Dieu, M., et al., Differential protein expression profiles in anterior gills of *Eriocheir sinensis* during acclimation to cadmium. *Aquat Toxicol* 2006, 76, 46-58.
- [57] Sethuraman, M., McComb, M. E., Huang, H., Huang, S., et al., Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *Journal of proteome research* 2004, 3, 1228-1233.
- [58] Dykens, J. A., Wiseman, R. W., Hardin, C. D., Preservation of phosphagen kinase function during transient hypoxia via enzyme abundance or resistance to oxidative inactivation. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology* 1996, 166, 359-368.
- [59] Fang, J. K., Wu, R. S., Chan, A. K., Yip, C. K., Shin, P. K., Influences of ammonia-nitrogen and dissolved oxygen on lysosomal integrity in green-lipped mussel *Perna viridis*: laboratory evaluation and field validation in Victoria Harbour, Hong Kong. *Mar Pollut Bull* 2008, 56, 2052-2058.
- [60] Malham, S. K., Cotter, E., O'Keeffe, S., Lynch, S., et al., Summer mortality of the Pacific oyster, *Crassostrea gigas*, in the Irish Sea: The influence of temperature and nutrients on health and survival. *Aquaculture* 2009, 287, 128-138.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.91

Identification of outer membrane proteins of *Edwardsiella tarda* in response to high concentration of copper

Chao Wang, Haili Zhang, Hui Li, Xuan-xian Peng

Center for Proteomics, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China

Received: 01 April 2012 Accepted: 27 July 2012 Available Online: 06 August 2012

ABSTRACT

The antimicrobial properties of copper have been reported, but little is known about the outer membrane proteins regulating copper resistance. In the present study, a sub-proteomic approach was utilized to investigate altered outer membrane (OM) proteins of *Edwardsiella tarda* in response to CuSO_4 . Upregulation of HemR and downregulation of Imp, EvpB, TolC, ETAE_2935, ETAE_1480, and ETAE_1723 were detected in *E. tarda* EIB202 survived in 1.0 mM CuSO_4 compared with the control without the ion. These alterations were validated, at random, using Western blotting. They were first revealed here to be bacterial copper-resistant proteins in combination of protein homology analysis. These findings highlight the way to clarify copper-resistant mechanisms.

Keywords: Copper resistance; Outer membrane proteins; *E. tarda*; Proteomics.

1. Introduction

Copper contributes to the function of numerous essential metabolic processes and thus is an essential micronutrient element required by almost living organisms [1]. However, copper at increased levels is also a unique metal known for its antimicrobial properties. The mechanism of copper ions toxic to cells has been extensively studied [2-4]. The most important antimicrobial mechanisms include: 1) Elevation of oxidative stress inside a cell result in oxidative damage to cells; 2) Decline in the membrane integrity of microbes leads to leakage of specific essential cell nutrients and subsequent cell death; 3) Unsuitable binding of copper to proteins that do not require copper for their function causes loss-of-function of the protein, and/or breakdown of the protein into nonfunctional portions.

The antimicrobial properties of copper are still under active investigation due to incompleteness of our knowledge to

antimicrobial role of copper. Recently, high throughput proteomics approach has been used to investigate the antimicrobial role of copper in *Escherichia coli*, *Edwardsiella tarda*, *Pseudomonas fluorescens*, *Pseudomonas putida*, *Lactococcus lactis* [5-9]. These studies elucidate the differential protein expression of these bacteria when exposed to copper stress, and indicate that the proteins from copper stressed cells exhibited a higher degree of oxidative proline and threonine modifications. However, information on alteration of OM proteins in response to copper stress is still absent.

Edwardsiella tarda is an uncommon enteric bacterium which has been found generally in animal hosts and occasionally in human feces [10]. Edwardsiellosis caused by the bacterium has been implicated in gastroenteritis, meningitis, biliary tract infections, peritonitis, liver and intra-abdominal abscesses, wound infections and septicemia [11, 12]. The

*Corresponding author: Dr. Xuanxian Peng, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, University City, Guangzhou 510006, People's Republic of China. Phone Number: +86-20-3145-2846; Fax Number: +86-20-8403-6215; E-mail Address: pxuanx@sysu.edu.cn; wangpeng@xmu.edu.cn

disease has caused severe economic losses in the fish farming industry worldwide. However, how to control the disease is still under investigation. A recent report has indicated that prolonged exposure to copper has multiple effects on *E. tarda* TX5 and results in significant attenuation of bacterial virulence [13]. In the present study, two-dimensional gel electrophoresis (2-DE) based proteomics was used to detect OM proteins of *E. tarda* EIB202 survived in medium with high concentration of CuSO_4 . Seven unique proteins were identified. Out of them, six were downregulated and one was up-regulated. These results may have significant implications in an understanding of mechanisms of *E. tarda* survived in high concentration of copper.

2. Material and Methods

2.1 Bacterial strain and culture

E. tarda EIB202, which complete genome sequence has been disclosed in 2009 [14], was kindly provided by Professor YX Zhang, East China University of Science and Technology, China. The bacterium was grown at 30°C in tryptic soy broth (TSB) medium overnight, and then diluted in 1:100 separately using the medium with or without 1.0 mM CuSO_4 . The cultures were harvested at an OD_{600} of 1.0 by centrifugation at 4,000g for 15min at 4°C.

2.2 Extraction of OM membrane proteins

The harvested bacterial cells were used for isolation of OM membrane proteins according to a procedure described previously [15]. Briefly, the cells were washed in sterile saline for three times, and then resuspended in sterile saline. Cells were disrupted by intermittent ultrasonic treatment. Unbroken cells and cellular debris were removed by centrifugation at 5,000g for 20 min and the supernatants were collected and were further centrifuged at 100,000g for 40 min at 4°C in a Beckman Coulter L-100XP centrifuge using a SW 41Ti Rotor. The precipitation was dissolved by 2% (W/V) sodium lauryl sarcosinate at room temperature for 30 min and ultracentrifuged again. The collecting pellets were resuspended in 50mM Tris-Cl and stored at -80°C. Concentrations of these proteins in the final preparation were determined using the Bradford method.

2.3 2-DE and MALDI-TOF analysis

2-DE was performed according to a procedure described previously [15]. Briefly, OM protein extracts containing 200 μg of proteins were dissolved in lysis solution and then rehydrate a pH 3-10 linear IPG strip (11 cm length, Bio-Rad). The total focusing was to be 40,000 Vhr and the maximum voltage was 8,000V using the Multiphor II system (Amersham). After the equalized by 2% (W/V) DTT and 2.5% (W/V) IAA (4M urea, 20% glycerol, 10% SDS) for 15

min, respectively, the IPG strips were transferred to the second-dimension electrophoresis using 12% acrylamide gels and stained with Coomassie Blue-R250. 2-DE gels were scanned in ImageScan and analyzed with ImageMaster 5.0 software (Amersham Biosciences, Sweden). Altered spots were standardized and then compared based on their volume percentages in the total spot volume over the whole gel image. Significantly changed spots were selected by rate increased / decreased ≥ 2 -fold or complete appearance and disappearance. Protein spots of interest were cut from the gels for analysis using Matrix-Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF). All MALDI analysis was performed by a fuzzy logic feedback control system (Reflex III MALDI-TOF system, Bruker) equipped with delayed ion extraction. Peptide mass finger printings were searched using the program Mascot (Matrix Science, London, U.K.) against the NCBI database; *E. coli* species database was defined as a matching species; one missed cleavage per peptide was allowed and the mass tolerance was 100 ppm. The protein subcellular locations were determined by Program PSORTb version 2.0 (<http://www.psort.org/psortb>).

2.4 Western blotting

Western blotting was performed as previously described [15]. Mouse antisera to TolC or EvpB were obtained from Fuji Bio Sci Tech Corp (Ji'an, China). Briefly, 1-DE and 2-DE gels were transferred to NC membranes for 3 h at 100 mA in transfer buffer (48 mM Tris, 39 mM glycine and 20% methanol) at 4 °C. The membranes were blocked with 5% skim milk and then incubated with mouse antisera to TolC or EvpB as the primary antibodies. The horseradish peroxidase (HRP) conjugated goat anti-mouse antibody was used as the secondary one. Antibody-tagged protein spots were detected by DAB.

2.5 Phylogenetic analysis

A BLAST search was performed on the sequences of these altered proteins using NCBI database as described previously [16]. A multiple sequence alignment was created with the amino acid sequences of YP_003294661.1 (ETA_E0603, Imp), YP_003296476.1 (ETA_E2430, EvpB), YP_003294249.1 (ETA_E0191, TolC), YP_003296979.1 (ETA_E2935), YP_003295532.1 (ETA_E1480), YP_003295773.1 (ETA_E1723), YP_003295845.1 (ETA_E1797, HemR) using ClustalX. Unrooted phylogenetic tree was constructed based on the sequences above using MEGA version 4.0 (<http://www.megasoftware.net>), and then was adjusted by TreeView to make it more readable. The sequences used in the trees were showed in Supplemental Figure 1. Bootstrap tests at 1000 replicates were carried out to examine the validity of the branching topologies. The numerical value of branch node reflects the confidence level. Values > 70 indicate statistical probability, while values < 50

are not statistically significant.

3. Results

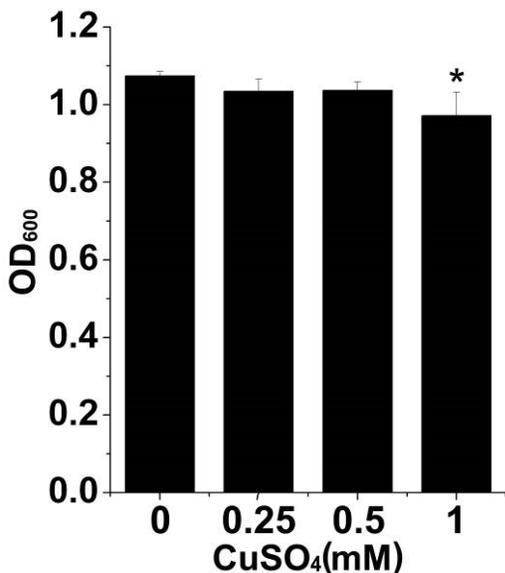


Figure 1. OD value of EIB202 cultured in medium with different concentrations of CuSO₄

3.1 Effect of CuSO₄ on EIB202 growth

EIB202 cells were cultured in TBS medium with 0, 0.25, 0.5 and 1.0 mM of CuSO₄. These cultures were harvested when the bacterial cell density in medium without the ions arrived at approximately 1.0 OD. Significantly inhibitory growth was found in bacteria cultured in medium with 1.0 mM CuSO₄ (Figure 1).

3.2 Identification of differential OM proteins of EIB202 in response to CuSO₄

To investigate an altered OM proteome of *E. tarda* in response to copper (II) ions, a sub-proteomic approach was utilized to identify differentially expressed proteins of sarcosine-insoluble fractions from EIB202 survived in medium with 1.0 mM of CuSO₄ (Cu-S). Approximately 40 protein spots were visualized in each of the gels stained with CBB R-250 (Figure 2A, B). Out of them, eight from Cu-S showed significant changes at the level of protein expression. They were named as 1 - 8. Figure 2C was an expanded view of the spots in the gels and Figure 2D showed the mean and standard deviation of these spots. They were identified as seven uniquely proteins. Table 1 showed the identities of these spots which were obtained by mass spectrometric analysis.

Table 1. Identification of Altered Spots by PMF or peptide fingerprinting Searching

Spot No.	Accession name	NCBI accession No.	Locus_tag	Protein description	Subcellular location	No. of peptides matched	No. of peptides unmatched	No. of peptides searched	Cover %	Mr/pI	NCBI score
1	D0ZC49_EDWTE	gi 269137961	ETA_E_06_03	organic solvent tolerance protein (Imp)	M/OM	50	40	90	52	90359/5.45	422
2	D0ZC49_EDWTE	gi 269137961	ETA_E_06_03	organic solvent tolerance protein (Imp)	M/OM	53	54	107	58	90359/5.45	435
3	D0ZB31_EDWTE	gi 269139775	ETA_E_24_30	type VI secretion system protein (EvpB)	unknown	33	54	87	57	54499/5.16	302
4	D0Z9U8_EDWTE	gi 269137549	ETA_E_01_91	outer membrane channel protein (TolC)	OM	17	21	38	49	51409/6.64	183
5	D0ZDY9_EDWTE	gi 269140278	ETA_E_29_35	nucleoside-specific channel-forming protein (Tsx)	OM	11	22	33	32	32786/5.62	116
6	D0ZH45_EDWTE	gi 269138831	ETA_E_14_80	MltA-interacting MipA family protein	unknown	17	74	91	39	31771/7.10	148
7	D0Z7Z3_EDWTE	gi 269139072	ETA_E_17_23	hypothetical protein ETA_E_1723	unknown	8	54	62	47	12695/7.67	112
8	D0Z865_EDWTE	gi 269139144	ETA_E_17_97	hemin receptor precursor (HemR)	OM/M	35	55	90	58	72809/6.29	302

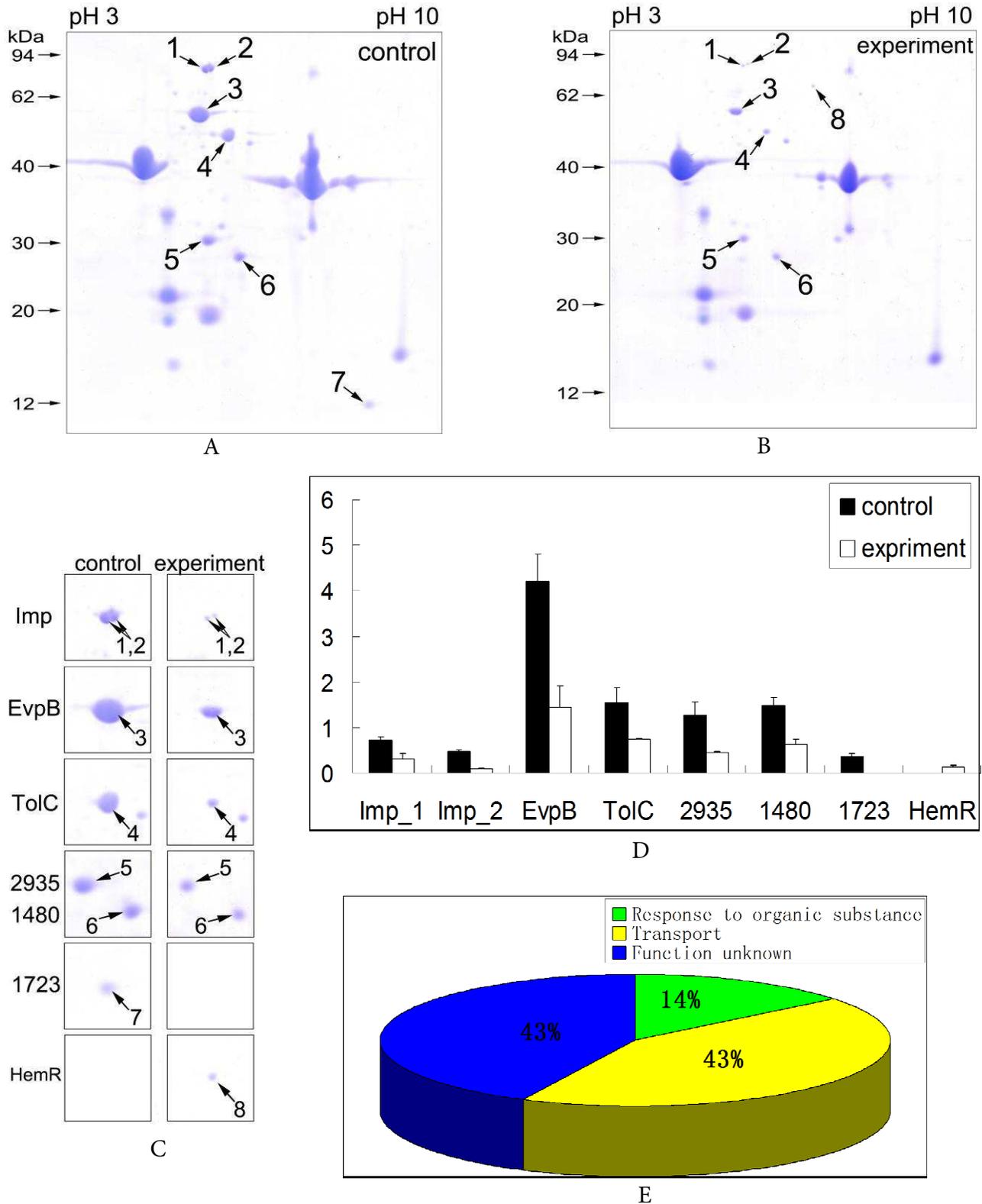


Figure 2. 2-DE subproteomics for investigation of altered OM proteins of *E. tarda* in response to copper. A, A representative 2-DE map of the sarcosine-insoluble fraction of Cu-C grown in TBS medium. B, A representative 2-DE map of the sarcosine-insoluble fraction of Cu-C grown in TBS medium. C, Enlarged partial 2-DE gels showing altered expression spots. D, Histogram displaying the changes in spot intensity of altered proteins between F-plasma-S-O (black) and F-plasma-S (white). Bars represent spot intensity, with the relative volume divided by the total volume over the whole image, according to ImageMaster version 5.0. E, Distribution of altered proteins based on GOA classification.

They were Imp (spot 1 and 2), EvpB (spot 3), TolC (spot 4), ETAE_2935 (spot 5), ETAE_1480 (spot 6), ETAE_1723 (spot 7), HemR (spot 8). Out of them, HemR was appeared and the others were decreased in Cu-S with respect to Cu-C. Except for EvpB, ETAE_1480 and ETAE_1723, which were proteins of unknown location, the other four were OM proteins. They are classified as response to organic substance (Imp), transport (TolC, ETAE_2935, HemR), function unknown (EvpB, ETAE_1480, ETAE_1723) (Figure 2E).

3.3 Validation of differential proteins using Western blotting analysis

To confirm these altered OM spots obtained from the 2-DE based proteomics, we randomly compare the difference of TolC and EvpB expression between Cu-C and Cu-S with 0.5, 1.0 mM concentrations of CuSO₄ using Western blotting. The Western blotting results indicated that downregulation of the two proteins was detected with the elevated concentrations of CuSO₄ (Figure 3). There was a better connection in TolC and EvpB between 2-DE and Western blotting. These results not only further demonstrated that TolC and EvpB were Cu-regulated expression proteins, but also indicated that there was a high confidence for the 2-DE based proteomics.

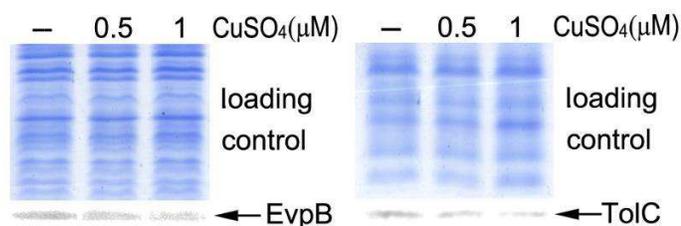


Figure 3. Confirmation of TolC and EvpB related to copper

3.4 Phylogenetic tree construction of altered proteins

Since information regarding to role of the altered proteins from *E. tarda* is largely absent, homology analysis was carried out to find the homology proteins from other bacteria for the functional analysis. Based on the amino acid sequences of the seven proteins from EIB202 deposited in GenBank, BLAST homology searching yielded the following closest matches. The unrooted phylogenetic analysis showed a close genetic relationship of the full-length proteins in EIB202 with more than ten Gram-negative bacteria including *Escherichia coli* (Suppl Figure 1). These results indicate that high homology between these EIB202 proteins and the proteins from other bacteria. We may understand role of the altered proteins in combination of the information of their homologies.

4. Discussion

In the present study, the OM proteome of *E. tarda* EIB202 in response to copper was identified. They were Imp, EvpB, TolC, ETAE_2935, ETAE_1480, and ETAE_1723. They belong to different functional classes of response to organic substance (Imp), transport (TolC, ETAE_2935, HemR), function unknown (EvpB, ETAE_1480, ETAE_1723) according to the EMBL-EBI GOA database (www.ebi.ac.uk/GOA). These proteins are first reported here to be involved in copper resistance.

Since the role of *E. tarda* proteins is largely unknown, we constructed the phylogenetic tree of these proteins from genetically close bacteria for understanding whether their homologies showed the role. Imp is one of organic solvent tolerance proteins that work for level of tolerance to organic solvents and the transport of lipopolysaccharide [17]. EvpB is one of type VI secretion system proteins that are widely distributed in pathogenic Gram-negative bacterial species [10]. TolC functions as a channel of the AcrAB-TolC multidrug efflux system. A report indicated that BaeSR, AcrD and MdtABC contribute to copper and zinc resistance in *Salmonella* [18]. Tsx is known to be involved in the permeation of nucleosides across the outer membrane under limiting substrate conditions and iron transport [19, 20]. MipS is an MltA-interacting MipA family protein. *E. coli* MipA serves as a scaffold protein required for the formation of a complex with MrcB/PonB and MltA and has recently been found to be involved in glucose regulation [19]. This complex could play a role in enlargement and septation of the murein sacculus and ETAE_1723 is a hypothetical protein. ETAE_1797 is a hemin receptor precursor. The receptor is involved in the utilization of heme and its protein complexes as iron sources in some of pathogenic bacteria [20]. In summary, TolC and hemin receptor are indirectly related copper metabolism, whereas the other five are first reported here to be copper-responsive proteins.

5. Concluding Remarks

In summary, a copper-responsive OM proteome is revealed in the present study. Out of the seven altered proteins, six were downregulated and four worked for transport. These results provide the proof on decline in the membrane integrity of microbes in response to high concentration of copper, but also show which OM proteins play the role in copper resistance.

6. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/91/0>

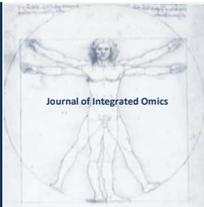
Supplementary material includes Figure 1. showing the unrooted phylogenetic analysis

Acknowledgements

This work was sponsored by grants from “863” project (2008AA092501), “973” project (2012CB114406) and NSFC projects (30972279, 40976080). Doctoral Fund of Ministry of Education of China (100171110029)

References

1. O.A. Karlse, O. Larse, H.B. Jense, H. Heipiepe. *FEMS Microbiol Lett* 323 (2011) 97-104.
2. Y. Ogr. *J Health Sci.* 57 (2011) 385-396
3. D.C. Kennedy, C.S. McKay, M.C.B. Legault, D. C. Danielson, J. A. J. A. Blake, A. F. Pegoraro, A. Stolow, Z. Mester, J. P. Pezacki. *J Am Chem Soc.* 133 (2011) 17993-18001
4. O. D. Iseri, D. A. Korpe, E. Yurtcu, F.I. Sahin, M. Habera. *Pant Cell Reports.* 30 (2011) 1713-1721
5. S. Sharma, C.S. Sundaram, P.M. Luthra, Y. Singh, R. Sirdeshmukh, W. N. Gad. *J Biotech* 126 (2006) 374-382
6. I. Poirier, N. Jean, J. C. Guary, M. Bertrand. *Sci Total Environ.* 406 (2008) 76-87
7. R. Nandakumar, C. E. Santo, N. Madayiputhiya, G. Grass. *Bio-metals.* 24 (2011) 429-444
8. C. D. Miller, B. Pettee, C. Zhang, M. Pabst, J.E. McLean, A.J. Anderso. *Lett Appl Microbiol.* 49 (2009) 775-783
9. O. Barre, F. Mourlane, M. Solioz. *J Bacteriol.* 189 (2007) 5947-5954
10. Y. Li, H. Zhan, Y. Li, H. L, X.X. Pen. *J Proteomics* 75 (2012) 1119 – 1128
11. H. Fan, X. J. Zhan, C. Z. Che, X. M. Ji, X. Y. Wan. *ACTA Oceanol Sin.* 25 (2006) 138-147
12. I. K. Wan, H. L. Ku, Y. M. Che, C. L. Li, H. Y. Chan, F. R. Chuan, M. H. Le. *Int J Clin Pract.* 59 (2005) 917-921
13. Y. H. Hu, W. Dang, C. S. Liu, L. Sun. *Lett Appl Microbiol.* 50 (2010) 97–103,
14. Q. Wan, M. Yan, J. Xia, H. W, X. Wan, Y. L, L. X, H. Zhen, S. Wan, G. Zha, Q. Li, Y. Zhan. *PLoS One.* 4 (2009) e7646.
15. C. Wan, Y. J. Li, H. L, W. J. X, H. Zhang, X. X. Pen. *J Proteomics* 2011. 75 (2012) 1263 – 1275
16. Z.H. Zhuang, X. L. Zhao, H. L, S. Y. Wang, X. X. Peng. *Fish Shellfish Immunol.* 31 (2011) 217-223
17. Bos, M (Bos, MP); Tefsen, (Tefsen, B); Geurtsen, (Geurtsen, J); Tommassen, . Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *PNAS.* 101 (2004) 9417-9422
18. K. Nishino, E. Nikaido, A. Yamaguch. Regulation of multidrug efflux systems involved in multidrug and metal resistance of *Salmonella enterica serovar typhimurium.* *J Bacteriol.* 189 (2007) 9066-9075



ORIGINAL ARTICLE | DOI: 10.5584/JIOMICS.V2I2.96

Comparative immunoproteome analysis of the response of susceptible A.BY/SnJ and resistant C57BL/6 mice to Coxsackievirus B3-infection

Elke Hammer^{*1}, Truong Quoc Phong¹, Leif Steil¹, Manuela Gesell Salazar¹, Christian Scharf¹, Reinhard Kandolf², Stephan B. Felix³, Heyo K. Kroemer⁴, Karin Klingel², Uwe Völker¹

¹Interfakultäres Institut für Genetik und Funktionelle Genomforschung, Universitätsmedizin Greifswald, Germany, F.-L.-Jahn-Str. 15^a, 17489 Greifswald; ²Abteilung Molekulare Pathologie, Universitätsklinikum Tübingen, Germany; ³Klinik für Innere Medizin B, Universitätsmedizin Greifswald, Germany; ⁴Institut für Pharmakologie, Universitätsmedizin Greifswald, Germany, Universitätsmedizin Greifswald, Germany.

Received: 26 June 2012 Accepted: 10 August 2012 Available Online: 11 August 2012

ABSTRACT

Both, innate and cell-mediated immunity contribute to prevention of chronic myocarditis and consecutively, cardiomyopathy. Thus, in resistant C57BL/6 mice myocarditis induced by Coxsackievirus B3 (CVB3)-infection is abrogated by immune-mediated mechanisms. However, susceptible A.BY/SnJ mice develop dilated cardiomyopathy (DCM) due to chronic myocarditis. Cardiac auto-antibodies have been shown to play a pivotal role in the initiation and/or progression of inflammatory DCM. In order to investigate differences in the autoimmune response of susceptible and resistant mice to infection with CVB3, the patterns of autoantibodies reacting with heart proteins in A.BY/SnJ and C57BL/6 mice were profiled by 2-D Western blot analysis during the acute and chronic phases of myocarditis up to three months, when the pathophysiological phenotype in the susceptible mice has progressed to DCM. In the early phase of infection both mouse strains displayed similar autoantibody patterns. In contrast, at later time points compared to the resistant C57BL/6 strain susceptible A.BY/SnJ mice displayed a much stronger autoimmune response against proteins associated with cell structure, protein transport as well as primary metabolic processes such as energy production. During chronic myocarditis strong antibody responses against myosin heavy chain 6, mitochondrial and heat shock proteins were observed in A.BY/SnJ mice. Antibodies directed against alpha-enolase, serotransferrin, radixin and two processed myosin protein species accumulated late and only in A.BY/SnJ mice suffering from inflammatory DCM. Functional assignment of the target proteins of cardiac autoantibodies indicates that these might be directly involved in cardiac dysfunction.

Keywords: Coxsackievirus B3-induced dilated cardiomyopathy; Murine model; Cardiac auto-antibodies; 2-D Western Blot.

1. Introduction

Dilated cardiomyopathy is a heart disease with an extremely poor clinical outcome. Besides genetic factors, mechanical stress and intoxication, virus infection-induced abnormalities in cellular and humoral immunity also contribute to the overall pathogenesis, such as cardiac enlargement and impaired systolic dysfunction.[1] Recent reports have provided evidence that autoimmune reactions against certain myocyte antigens may play a pivotal role in the initiation and/or progression of DCM. The pathogenic potential of cardiac auto-antibodies has been proven in animal models

either by active immunization with proteins like troponin I or by transfer of antibodies against the corresponding epitopes, both leading to dysfunction of the heart.[2-4] Furthermore, preliminary clinical data suggest that a few of these auto-antibodies are indeed “pathogenic”, actually causing cardiac dysfunction and heart failure.[5-6] Depending on the individual genetic pre-disposition such harmful autoimmune reactions are supposed to emerge as a consequence of heart muscle damage induced by viral triggers, ischemia or exposure to cardiotoxins leading to myocyte necrosis/

***Corresponding author:** Elke Hammer, Universitätsmedizin Greifswald, Interfakultäres Institut für Genetik und Funktionelle Genomforschung, Friedrich-Ludwig-Jahn-Str. 15^a, 17487 Greifswald. Phone: number: +49-3834-865872; Fax number: +49-3834-86795872; E-mail Address: hammer@uni-

apoptosis and subsequent liberation of a “critical amount” of self-antigens previously hidden to the immune system.[7-8]

Since extensive studies of human cardiac autoimmune reactions are limited by the restricted availability of appropriate tissue specimen such as endomyocardial biopsies, appropriate animal models are necessary. So far, only a few animal models that facilitate the analysis of pathogenesis of DCM following viral infection have been established.[9-11] Murine models of Coxsackievirus B3 (CVB3)-induced myocarditis excellently mimic the human disease patterns and are thus very well suited for the exploration of mechanisms leading to chronic myocarditis and finally to the dilatation of the heart. Development of severe myocarditis and synthesis of viral proteins during the acute phase of viral infection have been demonstrated in susceptible mouse strains such as A.BY/SnJ (H-2^b), DBA/2 (H-2^d) and SWR/J (H-2^a). In contrast, resistant C57BL/6 (H-2^b) and DBA/1J (H-2^a) mice eliminate the virus after an acute phase, do not show severe cardiac lesions and lack the chronic phase of the disease.[10, 12-15]

In order to investigate differences in the autoimmune response between susceptible and resistant mice infected with CVB3, the autoantibody patterns directed against heart proteins were followed by 2-D Western blot analysis over a period of 84 d post infection (p.i.) with CVB3 in A.BY/Sn and C57BL/6 mice. Over this time course of 3 months mice of the susceptible strain A.BY/SnJ displayed a much stronger autoimmune response than C57BL/6 mice. This response was primarily directed against proteins associated with cell structure, protein transport as well as primary metabolism. In chronic myocarditis strong antibody signatures against myosin heavy chain 6, mitochondrial and heat shock proteins were observed and their functional relevance is discussed.

2. Material and Methods

2.1 Virus infection and animal model of ongoing myocarditis

In this study we used a CVB3 stock which was derived from the infectious cDNA copy of the cardiotropic Nancy strain as previously described.[12] A.BY/SnJ mice (H-2^b) susceptible for the development of ongoing myocarditis and dilation of the heart following CVB3 infection and the resistant strain C57BL/6 were kept under specific pathogen-free conditions at the animal facilities of the Department of Molecular Pathology, University Hospital Tübingen. Experiments were conducted according to the German animal protection law. Mice (4-5 week-old, n=7 per time point) were infected intraperitoneally with 5×10^4 plaque-forming units (PFU) of purified CVB3 as described.[16] At different time points p.i., mice were sacrificed and hearts were collected for analysis. Histology and *in situ* hybridization was performed on transverse tissue sections covering the right and left ventricle taken below the level of the valves which were fixed in

4 % phosphate-buffered paraformaldehyde (pH 7.2) and embedded in paraffin. The rest of the left and right ventricles were snap frozen in liquid nitrogen and used for proteomic analysis.

2.2 Histology and *in situ* hybridization

Paraffin-embedded hearts were cut into 5 cm thick tissue sections and stained with hematoxylin/eosin (H&E) to evaluate the extent of myocardial injury and inflammation at day 0, 8, 28, and 84 p.i. For the detection of CVB3 positive-strand genomic RNA and interleukin-6 mRNA in hearts we used single-stranded ³⁵S-labeled RNA probes. Virus-specific probes were synthesized from the dual-promoter plasmid pCVB3-R1 as previously described.[12] Slide preparations were subjected to autoradiography, exposed for 3 weeks at 4 °C, and counterstained with H&E.

2.3 Preparation of protein extracts from heart tissues

Heart tissues were collected at four different time points representing different stages of disease: control – day 0, day 8 p.i., day 12 p.i., and day 84 p.i. Snap frozen heart tissue probes (n=2 per group) were disrupted to fine powder using a Mikro dismembrator (Braun, Melsungen, Germany) at 2.600 rpm for 2 min. The powder was dissolved in extraction buffer containing 8 M urea, 2 M thiourea, 2% CHAPS (Sigma, Taufkirchen, Germany) and sonicated on ice 3 times for 3 sec each with 9 cycles at 80 % energy using a Sonoplus (Bandelin, Berlin, Germany). Subsequently, the lysate was centrifuged at 16.000 g for 1 h at 4 °C and protein concentrations of the supernatants were determined using a Bradford assay kit (Bio-Rad, Munich, Germany). Sample aliquots were stored at -80 °C.

2.4 Preparation of immunoglobulins G (IgG) from murine sera

In order to estimate the dynamics of immune response to CVB3 infection in mice, pooled sera of seven mice were collected from the different time points (day 0, day 4 p.i., day 8 p.i., day 12 p.i., day 28 p.i., and day 84 p.i.) and used to isolate IgGs. Serum (250 µL) was mixed 1:1 (vol/vol) with Protein G Sepharose™ 4 Fast Flow (GE Healthcare, Munich, Germany) equilibrated with PBS buffer (138 mM NaCl, 2 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4, at 4 °C for 2 h) and incubated for 2 h with gentle shaking. After 5 washes with PBS buffer, the bound IgGs were eluted by adding 100 mM glycine, pH 2.5 for 15 min. The supernatant was collected and used to determine the amount of IgG using a Bradford assay kit (Bio-Rad, Munich, Germany).

2.5 2-D Western blot analysis

For 2-D Western blot analysis, a pool of heart proteins of the A.BY/SnJ strain from 4 different time points (0, 8, 12 and

84 days p.i.) was separated by 2-D gel electrophoresis using IPG strips of the pH-ranges 4-7 and 6-11, 11 cm length (GE Healthcare) and 12.5 % SDS-PAGE, and transferred onto PVDF membranes for Western blotting using a conventional semi-dry blotting device (Milliblot Graphic Electrobloetter II, Millipore, Billerica, MA, USA). For visualization of equal loading of protein, membranes were washed for 30 min in water and then stained with ink solution for 15 min before scanning of images with an Epson Expression 1680 Pro Scanner (Epson, Meerbusch, Germany). After destaining by soaking in TBS-T buffer (Tris-buffered saline-Tween, 20 mM Tris-HCl, 137 mM NaCl, and 0.1 % Tween 20), membranes were blocked with a solution of 5 % non-fat dry milk (Carl Roth, Karlsruhe, Germany) in TBS-T buffer for 90 min at room temperature. After 6 washes with TBS-T buffer, each membrane was incubated with a serum pool of seven individual mice from each time point as primary antibody mixture overnight at 4 °C. Membranes were washed 6 times in TBS-T buffer before the incubation with secondary antibody Immunopure peroxidase goat anti mouse IgG (H+L, 1:50.000) prepared in 5 % non-fat dry milk in TBS-T buffer at room temperature for 1 h. Following 6 washes in TBS-T buffer, membranes were incubated with SuperSignal West Femto Maximum Sensitivity Substrate (Pierce, Thermo Scientific, Bonn, Germany) for 5 min, before signals were detected with a Lumi-Imager (Roche, Mannheim, Germany).

2.6 Quantitative analysis of 2-D Western blots

Detection and quantification of spot signals on the Western blot images were performed using the Delta2D software version 3.4 (Decodon, Greifswald, Germany) with some adjustments. In order to suppress effects of the Western blot background, two parameters in the Delta2D software including “image filter closing” and “image filter opening” were set to 1 and 10, respectively. The mean absolute volume of 15 background spots evenly distributed on the blots of each time point was used for background subtraction. For further calculations only spots identified in both technical replicates were considered. The volumes of 6 different time points (day 0 and 4, 8, 12, 28, and 84 days p.i.) were used to evaluate the changes during the time course. The changes were determined by calculating the ratios of volumes of the corresponding time points divided by volumes at day 0.

2.7 Matrix assisted laser desorption ionization time-of-flight tandem mass spectrometry (MALDI-TOF-MS/MS)

In order to reveal which cardiac proteins were preferentially recognized by auto-antibodies following viral infection, spots of a 2-D gel loaded with 300 µg of protein and stained with Coomassie brilliant blue (Sigma, Taufkirchen, Germany) were matched with the corresponding Western blot signals via the corresponding protein map on the membranes recorded after ink staining before immunoblotting. Spots of interest were manually excised and subsequently subjected

to mass spectrometry for identification. Preparation of peptide extracts and MALDI-TOF-MS/MS analysis on a Proteome Analyzer 4800 (Applied Biosystems, Darmstadt, Germany) were carried out as described previously.[17] Identification of proteins was based on peptide mass fingerprint data confirmed by at least one protein specific peptide fragmentation pattern. Identifications with the International Protein Index (IPI) for mouse v. 3.26 using the Mascot algorithm via the GPS Explorer software package version 3.6 (Applied Biosystems) were considered to be statistically significant when the protein and peptide ion scores exceeded 60 and 32, respectively, which corresponds to $p < 0.05$. Methionine oxidation was set as variable and carbamidomethylation as fixed modification.

3. Results and discussion

3.1 Histopathological characterization of susceptible and resistant mice following CVB3 infection

Histopathological examination of A.BY/SnJ and C57BL/6 mouse hearts at different time points -0, 4, 8, 28, and 84 days p.i.- was performed to evaluate myocardial injury as well as virus infection and the disease state throughout the course of infection. Monitoring of the virus infection by radioactive *in situ* hybridization with ³⁵S-labeled CVB3-specific probes provided proof for strong focal viral replication in the acute phase of infection in both mice strains as already reported. [12, 18] Virus load was highest at day 8 p.i. in C57BL/6 mice, but declined very fast in the following days leading to complete virus clearance. Thus, no viral RNA was detectable in C57BL/6 mice at day 28 and 84 p.i.. In A.BY/SnJ mice virus load reached its maximum at day 8 p.i. as well. Subsequent to the acute phase of infection the intensity of virus replication dropped down, however, low amounts of viral RNA were still detectable at later stages of infection, documenting the well known CVB3 persistence in chronic myocarditis in A.BY/SnJ mice.[19-20] The magnitude of viral replication was paralleled in histological and immunohistological findings. At day 8 p.i., staining with H&E revealed that the extent of myocardial lesions and inflammation was higher in A.BY/SnJ than in C57BL/6 mice. Only in A.BY/SnJ mice ongoing inflammation and a significant fibrosis was noted which was followed by a dilated phenotype with a significant thinning of the ventricular walls at day 84 p.i. [21], whereas, no significant remodeling was observed in the resistant mouse strain. In C57BL/6 mice an early and adequate release of cytokines (interleukin (IL)-10, IL-6, tumor necrosis factor-alpha) by dendritic cells (DCs) was reported, which might contribute to prevention of chronic CVB3-induced myocarditis.[18] Additionally, the chemokines interferon-inducible protein 10 (IP-10) and chemokine (C-C motif) ligand 5 (RANTES) were found to be secreted by DCs from resistant C57BL/6 mice earlier in infection and at significantly higher levels, indicating a protective role of IP-10 and RANTES in

CVB3-induced myocarditis.[18]

3.2 Humoral immune response of susceptible and resistant mouse strain after CVB3-infection

In order to minimize the contribution of interindividual variations among individual mice, sera from seven different mice of each time point were pooled for the isolation of IgG and the determination of its total concentration. Figure 1 illustrates that the total IgG did not change dramatically immediately after infection (day 4 p.i.) compared to the concentration observed before infection. However, IgG level significantly increased afterwards, with A.BY/SnJ displaying a faster increase (day 8 p.i.) than C57BL/6 mice (day 12 p.i.). In both strains sera contained the highest amount of IgG antibodies at day 12 p.i. Thereafter, the amount of IgG decreased at day 28 p.i. in both mouse strains. However, the total amount of IgG was higher throughout the whole course of myocarditis in the A.BY/SnJ strain compared to C57BL/6 mice which is in contrast to the findings for CVB3-specific antibodies, where equal levels were observed in both mouse strains.[22] Moreover, in A.BY/SnJ mice the early response to virus infection was much stronger, displaying a 4.5-fold increase in IgG amount from day 4 to 8 p.i.. In C57BL/6 mice the IgG level increased only 2.9-fold in that time window, but at day 12 p.i. a more than 7-fold higher level compared to that before infection was reached. In summary, a lower IgG antibody level was observed in C57BL/6 mice prior to infection, which raised delayed compared to that of the A.BY/SnJ mouse strain and was not maintained at the same high level. Since the C57BL/6 mice started with a lower basal level, fold changes were larger following infection with CVB3 in this strain compared to A.BY/SnJ mice. Maintenance of higher levels of IgG in the A.BY/SnJ mice might represent the continuing interaction with CVB3 and the damaged host tissue.

3.3 Time-resolved analysis of the antibody patterns in A.BY/SnJ and C57BL/6 mice after infection with Coxsackievirus B3.

To explore possible differences in the autoimmune response of susceptible (A.BY/SnJ) and resistant (C57BL/6) mice at the different time points after infection with CVB3, the autoantibody repertoire was profiled using 2-D immunoblotting. The Western blot results document that applying regular standard limits of detection only very few weak signals were observed in pools of control sera of both mouse strains (Fig.2 A, B). Thus, laboratory animals kept under defined conditions did not seem to display any unspecific infections. In contrast, Western Blot signals strongly increased in intensity after infection with CVB3 from day 4 p.i. onwards (Fig.2 C, D, Supporting information Fig. S1).

For the permissive A.BY/SnJ mouse strain, a strong autoimmune response against cardiac autoantigens was observed up to day 28 p.i. (Fig. 2). Only at day 84 p.i. (Fig.2 K) a decrease in signal strength was recorded. For DCM patients fading of the cardiac auto-antibodies was noted with disease progression.[23]

In the resistant mouse strain C57BL/6, the strongest response was seen at day 4 p.i., followed by a continuous decrease of autoimmune response signals over time reflecting the reported virus elimination in these mice within 28 days p.i. [18] Importantly, significantly weaker signals were detected at later time points (day 28 and 84 p.i.), which was associated with complete recovery (Fig. 2 J, L). These different cardiac autoimmune response patterns and the differences in the total IgG patterns highlight different immune responses after viral infection in the two mouse strains. The observation of higher levels of IgG in A.BY/SnJ mice before infection fits with the hypothesis that higher frequencies of circulating auto-antibodies predispose subjects to develop more severe cardiac diseases like DCM. [24] This susceptibility to viral infection or immunization with heart-specific

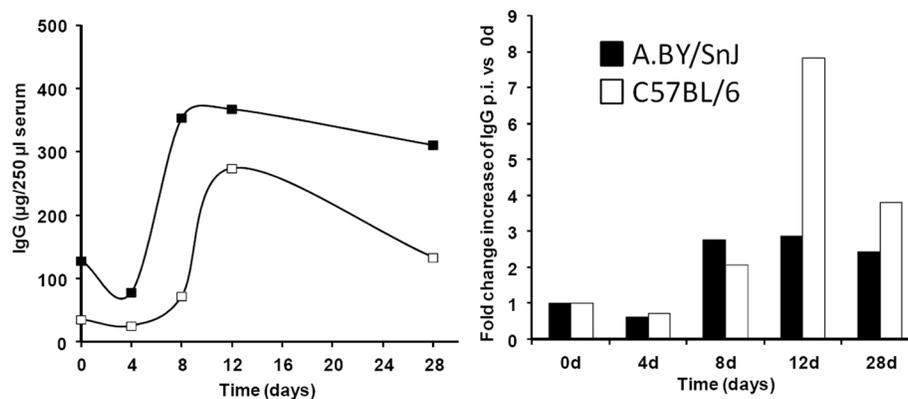


Fig. 1. Total IgG content in serum of A.BY/SnJ and C57BL/6 mice before and after infection with CVB3. A) Kinetics of changes of IgG concentrations during the time course of infection of a susceptible mouse strain A.BY/SnJ (white symbols) and a resistant mouse strain C57BL/6 (black). B) Relative changes in IgG-content compared to non-infected control samples.

auto-antigens of A.BY/SnJ mice is likely genetically determined [25] but the precise genetic alterations responsible have not been determined yet.

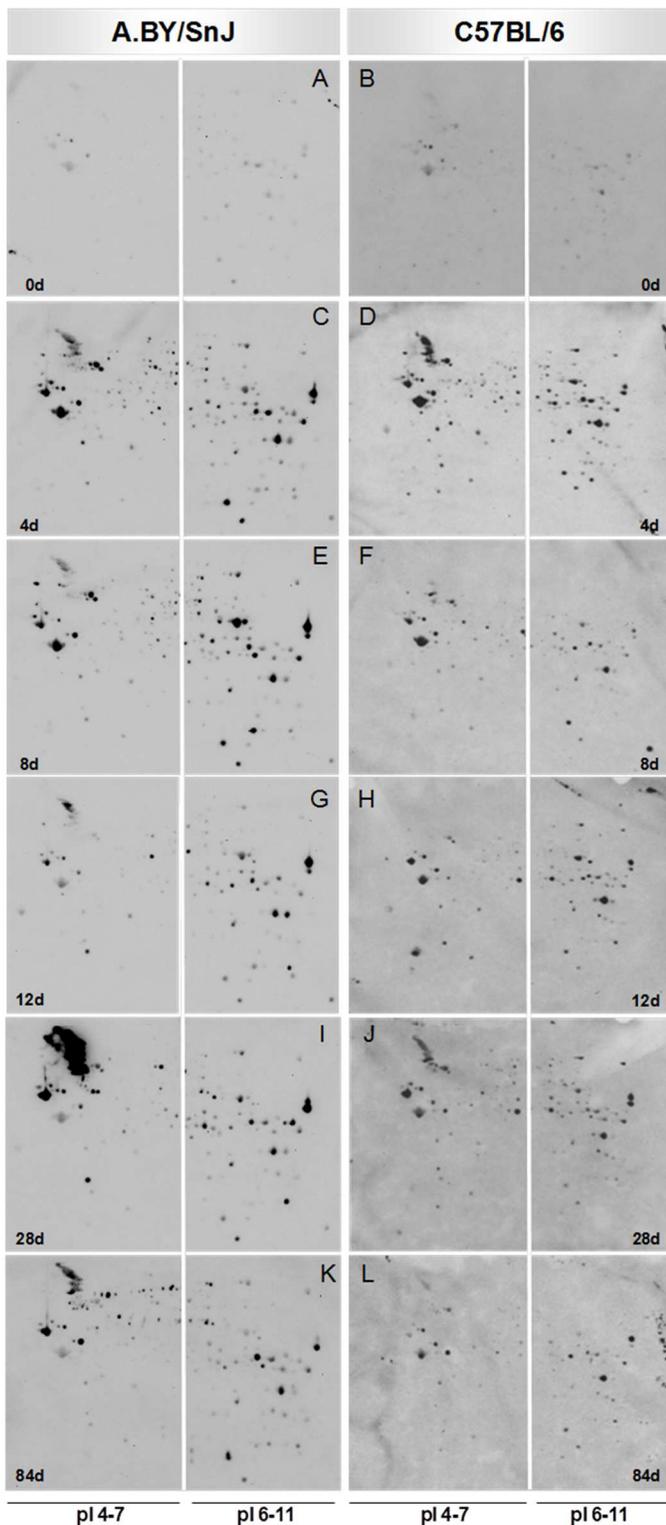


Fig. 2. Kinetics of autoimmune response of A.BY/SnJ and C57BL/6 mice against Coxsackievirus infection. 2D Western Blot images using IPG strips pH 4-7 and pH 6-11 were combined and display the autoimmune responses in both mice strains compared to non-infected controls (day 0).

A quantitative analysis of Western blot signals with the aid of the Delta2D software revealed a total of 227 spot signals for A.BY/SnJ mice, while only 169 spot signals were detected for the C57BL/6 strain (Supporting information Table S1). 131 spot signals were observed in the autoantibody patterns of both strains, whereas, 96 and 38 spot signals were only present in A.BY/SnJ and C57BL/6 mice, respectively. Due to the very low intensities of antibody signals observed in the non-infected mice, for further quantitative analysis only spot signals exceeding a fold increase >5 in the intensity ratio between the particular time point and the non-infected control group were considered. The number of significantly increased signals was higher in A.BY/SnJ mice compared to C57BL/6 mice during the whole course of infection (Supporting information Table S1). The autoimmune response due to viral infection was highest at day 4 p.i. in both mouse strains with 82 % (A.BY/SnJ) and 69 % (C57BL/6) of all antibody signals exceeding 5-fold increase. Remarkably, this peak in the cardiac autoimmune response clearly preceded the increase in total IgG level in both strains (see Fig. 1). However, while the level of most antibodies declined rapidly in the resistant strain, in A.BY/SnJ mice about 50 % of all antibodies displayed such high levels throughout the chronic phase up to day 84 p.i. (Fig. 3). High titers of auto-antibodies can influence cardiac function by negative chronotropic and/ or negative inotropic effects mediated by binding to surface receptors, forming immune complexes or influencing downstream processes. [26]

3.4 Identification of protein targets for auto-antibodies

Identification of antigens by mass spectrometric analysis revealed protein identities for 209 of 227 spots visible on the Coomassie-stained gels (92 %) representing 122 distinct protein species.[27] For A.BY/SnJ mice 8 spots contained 2 proteins, and the remaining 201 spots contained only one major protein (Supporting information Fig. S2, Table S2). Similarly for strain C57BL/6, proteins have been identified in 125 of 169 spots (74 %) representing 124 proteins (87 distinct species). Here, only one spot contained 2 proteins, and the remaining 124 spots each contained only one predominant protein (Supporting information Table S3). Thus, the identifications of antibody targets by MALDI-mass spectrometry indeed permitted an evaluation of the protein groups targeted. The majority of proteins representing antigens in both mouse strains were of mitochondrial origin (44 %), followed by cytoplasmic (19 %), cytoskeletal (7 %) and plasma membrane proteins (7 %). The low coverage of surface antigens was likely caused by the fact that proteins with multiple transmembrane domains cannot be resolved efficiently on 2-D gels.

Although antibody formation against intracellular proteins is not expected, high abundance of heart specific mitochondria targeting antibodies (M7) was already reported in DCM patients in 1984.[28] The functional relevance of antibody production was also shown in a CVB3 infection model

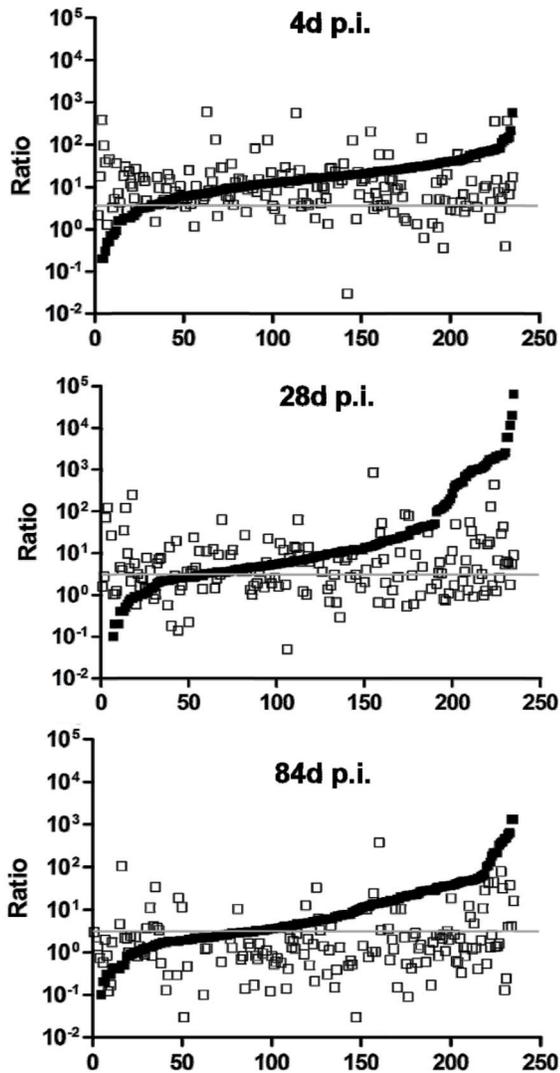


Fig. 3. Kinetics of Western Blot signal intensities at 4, 28 and 84 days p.i. with CVB3 in comparison to non-infected mice (day 0). Ratios of each time point vs. day 0 of A.BY/SnJ mice (black) were sorted by size and the corresponding ratio of each signal in C57BL/6 mice (white) was plotted. The horizontal gray line marks a 5 fold increase of the antibody signal in comparison to non-infected mice.

where a strong increase of antibodies directed against the mitochondrial adenine-nucleotide transporter led to alterations in cellular energy consumption and calcium homeostasis. [29, 30] The localization of the antigens for which antibody signals were identified in A.BY/SnJ mice only (48 protein species) did not differ significantly from that of C57BL/6 mice except for a slightly increased percentage of cytoskeletal antigens (10 %).

Functional categorization of all identified proteins (115 different protein species mapped to Uniprot accession numbers) disclosed carbohydrate metabolism (30 proteins), respiratory electron transport chain (19), cellular component organization (12) and tricarboxylic acid cycle (9) as the major biological processes targeted by auto-reactive antibodies (Supporting information Table S4). In the acute phase of

infection (day 4 and 8 p.i.) the antigens targeted by the cardiac autoantibodies were assigned to similar biological processes in both mouse strains (according to PANTHER-www.pantherdb.org). However, a comparison of the proteins recognized by autoantibodies in A.BY/SnJ and C57BL/6 mice at particular time points revealed a higher number and stronger signal intensities in the permissive strain for targets involved in primary metabolic processes like respiration, carbohydrate metabolism and tricarboxylic acid cycle but

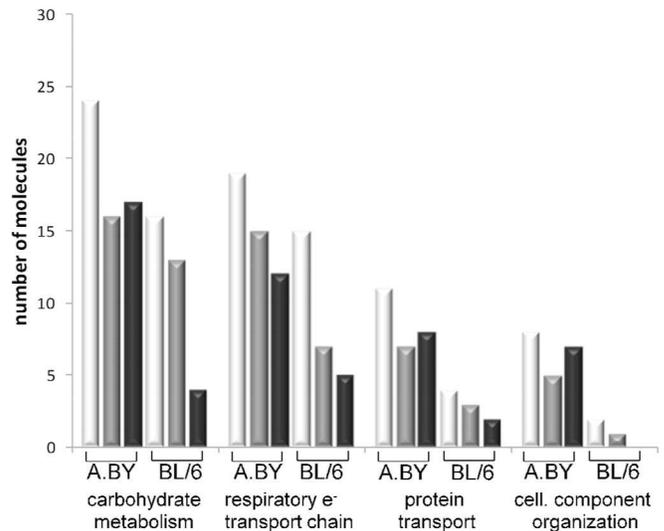


Fig. 4. Biological processes displaying differences in the number of assigned autoantibodies (assigned by the identification of the corresponding antigen) between the susceptible strain A.BY/SnJ and the resistant strain C57BL/6 at day 84 p.i. (black bars). Acute myocarditis (day 4 and day 8 p.i.) white bars; chronic myocarditis (day 28 p.i. grey bars).

also protein transport and cellular component organization (Supporting information Table S5). This higher number of protein targets in the susceptible strain A.BY/SnJ (Fig. 4, Supporting information Table S1 and S5) might reflect more pronounced cardiomyocyte lysis even in the early phase of infection.[31]

In agreement with observations for various animal models [5, 32] myosin heavy chain 6 was the most pronounced cardiac antigen in the permissive strain for which signal intensities increased up to 1800-fold at day 28 p.i. compared to the non-infected controls. A similar extremely strong signal was also reported for CVB3 infected SWR/Ola (H-2q) mice at day 25 p.i. suffering from subsequent myocarditis but without mortality.[33] However, rather weak anti-myosin antibody signals were observed in C57BL/6 mice at day 28 p.i., Molecular mimicry between Coxsackievirus and myosin was hypothesized to play a role in pathogenesis.[34-35] However, anti-myosin antibodies were also detected after troponin I-induced myocardial damage [36] or immunization with myosin itself.[37] These reports as well as the marked production of anti-myosin antibody in the permissive strain support a causative role of released cardiac antigens and their

corresponding auto-antibodies in pathogenesis of cardiomyopathies.[1] Auto-antibodies against troponin I or T, which were reported in other studies (reviewed by [38]), could not be detected in this study although the respective antigens were abundantly present in the 2-DE (data not shown). These differences might be caused by different genetic background because analyses of human sera from patients suffering from DCM and ICM against known cardiac auto-antibodies –among them anti-troponin I- revealed only frequencies of 15-20 % presence of specific antibodies [39] supporting the notion of a highly variable autoimmune response following myocardial damage. In our study we cannot address the prevalence of auto-antibodies directed against cardiac G-protein coupled receptors [40-41] since these membrane proteins were not resolved on the 2-D-gels [42] used for pre-fractionation of proteins prior to immunoblotting.

In A.BY/SnJ mice an autoantibody profile very similar to the one just described for myosin heavy chain 6 auto-antibodies, i.e. highest intensities at day 28 p.i. followed by a substantial drop in titer during active DCM (day 84 p.i.), was also observed for a number of mitochondrial proteins like components of complex I, II, III, and V of the respiratory electron transfer chain as well as for acyl-CoA dehydrogenases. A variety of these proteins are flavoproteins (medium-chain specific acyl-CoA dehydrogenase, long-chain specific acyl-CoA dehydrogenase, isovaleryl-CoA dehydrogenase, electron transfer flavoprotein subunit beta, succinate dehydrogenase flavoprotein subunit Sdha and [ubiquinone] iron-sulfur protein Sdhb) which have already been described as cardiac antigens of auto-antibodies after CVB3-infection in mice.[43] Since mitochondria are highly abundant in cardiomyocytes, protein release in case of myocardial damage might lead to substantial antibody formation with functional relevance. The high number of gap-junctions facilitating the penetration of plasma membrane allows antibodies to gain access even to intracellular targets [30] and might contribute to energy limitations supposed already from altered levels of transcripts and proteins involved in lipid metabolism [21] and respiratory electron transport chain.[44-45]

Also for stress proteins like mitochondrial stress-70 protein and heat shock protein 8 strong increases in Western Blot signal intensities were detected predominantly at day 28 p.i. in comparison to non-infected mice. Cardiac autoantibodies against heat shock proteins have also been found in serum of patients suffering from DCM [46], supporting the similarities in the autoimmunity-related processes observed in animal models and in patients.

With the manifestation of the cardiomyopathy phenotype a drop in cardiac autoantibody levels was detected, although the titers in the permissive mice remained at much higher level than those in the recovered strain C57BL/6 (Figure 2I, J, Fig. 5). Most of the intense signals decreased significantly up to day 84 p.i. as already mentioned.

However, for a small group of cardiac proteins like radixin, alpha enolase, sero transferrin and 2 myosin peptides

high antibody titers were detected only in the susceptible mouse strain and only at 84 days p.i. (Supporting information Table S1). In contrast, no antibodies against these proteins were present in C57BL/6 mice at any time point of myocarditis. Thus, a specific role for these auto-antibodies in virus-induced cardiomyopathy can be speculated, but has to be proven by use of age matched controls and the analysis of sera from different animal models as well patient samples.

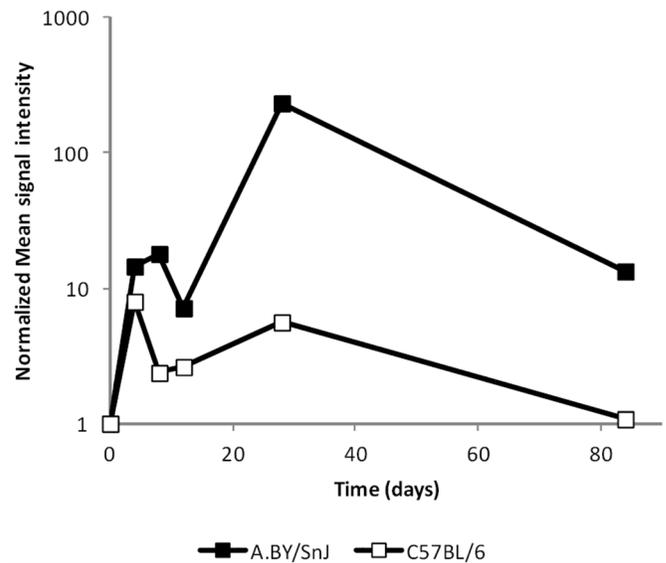


Fig. 5. Kinetics of mean Western Blot signal intensities in comparison to non-infected mice (day 0) of A.BY/SnJ mice (black symbols) and C57BL/6 mice (white symbols).

4. Concluding Remarks

In this study we investigated two mouse strains differing in their immune response after infection with CVB3, thus displaying a different cardiac autoantibody profile and repertoire. In the resistant C57BL/6 strain which is able to eliminate the virus within a few days due to effective cytokine production and activation of immune cells only short term accumulation of cardiac auto-antibodies was observed, the titer of which then declined fast and significantly. In contrast, in the permissive strain A.BY/SnJ virus persistence and ongoing myocardial injury were associated with persistent high titers of cardiac auto-antibodies. Here, a certain genetic predisposition probably causes alterations in the adaptive immune system. Furthermore, linkage of the major histocompatibility complex to susceptibility to chronic viral infections has been hypothesized. More extensive comparative genetic studies of mice strains presenting extensive autoimmune reactions with those lacking strong auto-antibody formation after viral infection are necessary, and the two strains of the present study seem to be potential candidates for such studies. The time resolved profile of the autoimmune response of A.BY/SnJ mice was characterized by the presence of massive antibody titers against cardiac myosin and mitochondrial proteins during chronic inflammation

(day 28 p.i.). The supposed lack of energy due to inference of antibodies with primary metabolism as well as loss of sarcomer integrity due to targeting of cell structure proteins might contribute to further myocyte loss leading later on to the manifestation of cardiomyopathy as a late consequence of enterovirus infection.

5. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/96/0>.

Supplementary material includes:

- Supplementary Fig. S1. Number of cardiac autoantibodies detected in serum of mice infected with CVB3.
- Supplementary Fig. S2 Fused images of 2-D Western Blot images of all time points of pooled serum.
 - * A. Labeled (spot signal numbers) fused image of 2-D Western Blot images of all time points of pooled serum in the pH range 4-7.
 - * B. Labeled (Gene name or UniProt accession number) fused image of 2-D Western Blot images of all time points of pooled serum in the pH range 4-7.
 - * C. Labeled (spot signal numbers) fused image of 2-D Western Blot images of all time points of pooled serum in the pH range 6-11.
 - * D. Labeled (Gene name or UniProt accession number) fused image of 2-D Western Blot images of all time points of pooled serum in the pH range 6-11.
- Supplementary Table S1. List of all autoantigens identified by MALDI-TOF-MS/MS corresponding to autoantibodies generated in two mouse strains.
- Supplementary Table S2. Protein identification details of cardiac protein targets of autoantibodies detected by 2-D Western Blot in serum of A.BY/SnJ mice.
- Supplementary Table S3. Protein identification details of cardiac protein targets of autoantibodies detected by 2-D Western Blot in serum of A.BY/SnJ mice.
- Supplementary Table S4. List of biological processes displaying significant enrichment ($p < 0.05$) among the list of proteins identified as antigens for autoantibodies in the hearts of A.BY/SnJ or C57BL/6 mice infected with CVB3. Enrichment of biological processes among the regulated proteins was calculated in the PANTHER software by comparison with the reference set of NCBI mouse database.
- Supplementary Table S5. List of biological processes displaying significant enrichment ($p < 0.05$) among the list of proteins identified as antigens for autoantibodies in the hearts of A.BY/SnJ or C57BL/6 mice infected with CVB3 resolved per time point. Enrichment of biological processes among the regulated proteins was calculated in the PANTHER software by comparison with the reference set of NCBI mouse database. Color defines significance of enrichment: dark red: highly significant; bright red $> 0.01 < 0.05$; yellow no significant enrichment.

Acknowledgements

We are grateful to Decodon GmbH, Greifswald, Germany for providing unlimited access to pre-release versions of their Delta-2D software package. The study was performed

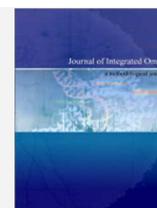
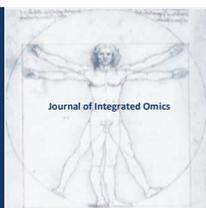
within the framework of the collaborative research project SFB/TR19 and supported by grants of the Deutsche Forschungsgemeinschaft to S.B.F., K.K. and U.V.

References

1. T. Yoshikawa, A. Baba, Y. Nagatomo, Autoimmune mechanisms underlying dilated cardiomyopathy. *Circ. J.* 73 (2009) 602-607. DOI: JST.JSTAGE/circj/CJ-08-1151.
2. R. Jahns, V. Boivin, L. Hein, S. Triebel, C. E. Angermann, G. Ertl, M. J. Lohse, Direct evidence for a beta 1-adrenergic receptor-directed autoimmune attack as a cause of idiopathic dilated cardiomyopathy. *J. Clin. Invest.* 113 (2004) 1419-1429. DOI: 10.1172/JCI20149.
3. T. Okazaki, Y. Tanaka, R. Nishio, T. Mitsuiye, A. Mizoguchi, J. Wang, M. Ishida, H. Hiai, A. Matsumori, N. Minato, T. Honjo, Autoantibodies against cardiac troponin I are responsible for dilated cardiomyopathy in PD-1-deficient mice. *Nature Med.* 9 (2003) 1477-1483. DOI: 10.1038/nm955.
4. R. S. Warraich, E. Griffiths, A. Falconar, V. Pabbathi, C. Bell, G. Angelini, M. S. Suleiman, M. H. Yacoub, Human cardiac myosin autoantibodies impair myocyte contractility: a cause-and-effect relationship. *FASEB J* 20 (2006) 651-660. DOI: 10.1096/fj.04-3001.com.
5. L. Caforio, M. Grazzini, J. M. Mann, P. J. Keeling, G. F. Bottazzo, W. J. McKenna, S. Schiaffino, Identification of alpha- and beta-cardiac myosin heavy chain isoforms as major autoantigens in dilated cardiomyopathy. *Circulation* 85 (1992) 1734-1742. DOI: 10.1161/01.CIR.85.5.1734.
6. R. Jahns, V. Boivin, V. Schwarzbach, G. Ertl, M. J. Lohse, Pathological autoantibodies in cardiomyopathy. *Autoimmunity* 41 (2008) 454-461. DOI: 10.1080/08916930802031603.
7. U. Eriksson, R. Ricci, L. Hunziker, M. O. Kurrer, G. Y. Oudit, T. H. Watts, I. Sonderegger, K. Bachmaier, M. Kopf, J. M. Penninger, Dendritic cell-induced autoimmune heart failure requires cooperation between adaptive and innate immunity. *Nature Med.* 9 (2003) 1484-1490. DOI: 10.1038/nm960.
8. S. A. Huber, P. A. Lodge, Coxsackievirus B-3 myocarditis in Balb/c mice. Evidence for autoimmunity to myocyte antigens. *Am. J. Pathol.* 116 (1984) 21-29.
9. D. Cihakova, N. R. Rose, Pathogenesis of myocarditis and dilated cardiomyopathy. *Adv. Immunol.* 99 (2008) 95-114. DOI: 10.1016/S0065-2776(08)00604-4.
10. D. Fairweather, N. R. Rose, Coxsackievirus-induced myocarditis in mice: a model of autoimmune disease for studying immunotoxicity. *Methods* 41 (2007) 118-122. DOI: 10.1016/j.jymeth.2006.07.009.
11. R. Wessely, Coxsackieviral replication and pathogenicity: lessons from gene modified animal models. *Med. Microbiol. Immunol.* 193 (2004) 71-74. DOI: 10.1007/s00430-003-0203-0.
12. K. Klingel, C. Hohenadl, A. Canu, M. Albrecht, M. Seemann, G. Mall, R. Kandolf, Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue damage, and inflammation. *Proc. Natl. Acad. Sci. U. S. A.* 89 (1992) 314-318.
13. C. Leipner, K. Grun, I. Schneider, B. Gluck, H. H. Sigusch, A. Stelzner, Coxsackievirus B3-induced myocarditis: differences in the immune response of C57BL/6 and Balb/c mice. *Med. Microbiol. Immunol.* 193 (2004) 141-147. DOI: 10.1007/s00430-003-0199-5.
14. Matsumori, C. Kawai, An animal model of congestive (dilated) cardiomyopathy: dilatation and hypertrophy of the

- heart in the chronic stage in DBA/2 mice with myocarditis caused by encephalomyocarditis virus. *Circulation* 66 (1982) 355-360. DOI: 10.1161/01.CIR.66.2.355.
15. L. J. Wolfgram, K. W. Beisel, A. Herskowitz, N. R. Rose, Variations in the susceptibility to Coxsackievirus B3-induced myocarditis among different strains of mice. *J. Immunol.* 136 (1986) 1846-1852. DOI: 0022-1767/86/1365-1846\$0200/0.
 16. R. Kandolf, H. C. Selinka, K. Klingel, in: B. L. Semler, E. Wimmer (Eds.), *Molecular Biology of Picornaviruses*, American Society for Microbiology, Washington, DC 2002, pp. 405-413.
 17. T. Thiele, L. Steil, S. Gebhard, C. Scharf, E. Hammer, M. Brigulla, N. Lubenow, K. J. Clemetson, U. Volker, A. Greinacher, Profiling of alterations in platelet proteins during storage of platelet concentrates. *Transfusion* 47 (2007) 1221-1233. DOI: 10.1111/j.1537-2995.2007.01255.x.
 18. O. Weinzierl, G. Szalay, H. Wolburg, M. Sauter, H. G. Ramnensee, R. Kandolf, S. Stevanovic, K. Klingel, Effective chemokine secretion by dendritic cells and expansion of cross-presenting CD4-/CD8+ dendritic cells define a protective phenotype in the mouse model of coxsackievirus myocarditis. *J. Virol.* 82 (2008) 8149-8160. DOI:10.1128/JVI.00047-08.
 19. M. Esfandiarei, B. M. McManus, Molecular biology and pathogenesis of viral myocarditis. *Annu Rev Pathol* 3 (2008) 127-155. DOI: 10.1146/annurev.pathmechdis.3.121806.151534.
 20. M. Bonneau, A. Darveau, N. Sonenberg, Effect of viral infection on host protein synthesis and mRNA association with the cytoplasmic cytoskeletal structure. *J. Cell Biol.* 100 (1985) 1209-1218. DOI: 0021-9525/85/04/1209/10 \$1.00
 21. K. Nishtala, T. Q. Phong, L. Steil, M. Sauter, M. G. Salazar, R. Kandolf, H. K. Kroemer, S. B. Felix, U. Volker, K. Klingel, E. Hammer, Virus-induced dilated cardiomyopathy is characterized by increased levels of fibrotic extracellular matrix proteins and reduced amounts of energy-producing enzymes. *Proteomics* 11 (2011) 4310-4320. DOI: 10.1002/pmic.201100229.
 22. G. Szalay, S. Meiners, A. Voigt, J. Lauber, C. Spieth, N. Speer, M. Sauter, U. Kuckelkorn, A. Zell, K. Klingel, K. Stangl, R. Kandolf, Ongoing coxsackievirus myocarditis is associated with increased formation and activity of myocardial immunoproteasomes. *Am. J. Pathol.* 168 (2006) 1542-1552. DOI: 10.2353/ajpath.2006.050865.
 23. L. Caforio, J. H. Goldman, M. K. Baig, A. J. Haven, L. Dalla Libera, P. J. Keeling, W. J. McKenna, Cardiac autoantibodies in dilated cardiomyopathy become undetectable with disease progression. *Heart* 77 (1997) 62-67.
 24. L. Caforio, A. Vinci, S. Iliceto, Anti-heart autoantibodies in familial dilated cardiomyopathy. *Autoimmunity* 41 (2008) 462-469. DOI:10.1080/08916930802031546.
 25. F. Leuschner, H. A. Katus, Z. Kaya, Autoimmune myocarditis: past, present and future. *J. Autoimmun.* 33 (2009) 282-289. DOI: 10.1016/j.jaut.2009.07.009.
 26. Z. Kaya, C. Leib, H. A. Katus, Autoantibodies in heart failure and cardiac dysfunction. *Circ. Res.* 110 (2012) 145-158. DOI: 10.1161/CIRCRESAHA.111.243360.
 27. P. R. Jungblut, H. G. Holzthutter, R. Apweiler, H. Schluter, The speciation of the proteome. *Chem. Cent. J.* 2 (2008) 16. DOI: 10.1186/1752-153X-2-16.
 28. R. Klein, B. Maisch, K. Kochsiek, P. A. Berg, Demonstration of organ specific antibodies against heart mitochondria (anti-M7) in sera from patients with some forms of heart diseases. *Clin. Exp. Immunol.* 58 (1984) 283-292.
 29. K. Schulze, B. F. Becker, R. Schauer, H. P. Schultheiss, Antibodies to ADP-ATP carrier--an autoantigen in myocarditis and dilated cardiomyopathy--impair cardiac function. *Circulation* 81 (1990) 959-969. DOI: 10.1161/01.CIR.81.3.959.
 30. K. Schulze, B. Witzendichler, C. Christmann, H. P. Schultheiss, Disturbance of myocardial energy metabolism in experimental virus myocarditis by antibodies against the adenine nucleotide translocator. *Cardiovasc. Res.* 44 (1999) 91-100. DOI: S0008-6363(99)00204-7.
 31. G. Szalay, M. Sauter, M. Haberland, U. Zuegel, A. Steinmeyer, R. Kandolf, K. Klingel, Osteopontin: a fibrosis-related marker molecule in cardiac remodeling of enterovirus myocarditis in the susceptible host. *Circ. Res.* 104 (2009) 851-859. DOI: 10.1161/CIRCRESAHA.109.193805.
 32. N. R. Rose, A. Herskowitz, D. A. Neumann, N. Neu, Autoimmune myocarditis: a paradigm of post-infection autoimmune disease. *Immunol. Today* 9 (1988) 117-120. DOI: 10.1016/0167-5699(88)91282-0.
 33. N. Latif, H. Zhang, L. C. Archard, M. H. Yacoub, M. J. Dunn, Characterization of anti-heart antibodies in mice after infection with coxsackie B3 virus. *Clin. Immunol.* 91 (1999) 90-98. DOI: 10.1006/clim.1998.4679.
 34. M. W. Cunningham, Cardiac myosin and the TH1/TH2 paradigm in autoimmune myocarditis. *Am. J. Pathol.* 159 (2001) 5-12. DOI: 10.1016/S0002-9440(10)61665-3.
 35. J. Gauntt, P. T. Gomez, P. S. Duffey, J. A. Grant, D. W. Trent, S. M. Witherspoon, R. E. Paque, Characterization and myocarditic capabilities of coxsackievirus B3 variants in selected mouse strains. *J. Virol.* 52 (1984) 598-605. DOI: 0022-538X/84/110598-08\$02.00/0.
 36. S. Goser, M. Andrassy, S. J. Buss, F. Leuschner, C. H. Volz, R. Ottl, S. Zittrich, N. Blaudeck, S. E. Hardt, G. Pfitzer, N. R. Rose, H. A. Katus, Z. Kaya, Cardiac troponin I but not cardiac troponin T induces severe autoimmune inflammation in the myocardium. *Circulation* 114 (2006) 1693-1702. DOI:10.1161/CIRCULATIONAHA.106.635664.
 37. N. Neu, N. R. Rose, K. W. Beisel, A. Herskowitz, G. Gurri-Glass, S. W. Craig, Cardiac myosin induces myocarditis in genetically predisposed mice. *J. Immunol.* 139 (1987) 3630-3636. DOI: 0022-1767/87/1311-3630\$02.00/0
 38. Z. Kaya, H. A. Katus, N. R. Rose, Cardiac troponins and autoimmunity: their role in the pathogenesis of myocarditis and of heart failure. *Clin. Immunol.* 134 (2010) 80-88. DOI: 10.1016/j.clim.2009.04.008.
 39. M. Landsberger, A. Staudt, S. Choudhury, C. Trimpert, L. R. Herda, K. Klingel, R. Kandolf, H. P. Schultheiss, H. K. Kroemer, U. Volker, S. B. Felix, Potential role of antibodies against cardiac Kv channel-interacting protein 2 in dilated cardiomyopathy. *Am. Heart J.* 156 (2008) 92-99. DOI: j.ahj.2008.02.015.
 40. R. Jahns, V. Boivin, T. Krapf, G. Wallukat, F. Boege, M. J. Lohse, Modulation of beta1-adrenoceptor activity by domain-specific antibodies and heart failure-associated autoantibodies. *J. Am. Coll. Cardiol.* 36 (2000) 1280-1287. DOI: S0735-1097(00)00881-0.
 41. Baba, T. Yoshikawa, Y. Fukuda, T. Sugiyama, M. Shimada, M. Akaishi, K. Tsuchimoto, S. Ogawa, M. Fu, Autoantibodies against M2-muscarinic acetylcholine receptors: new upstream targets in atrial fibrillation in patients with dilated cardiomyopathy. *Eur. Heart J.* 25 (2004) 1108-1115. DOI: j.ehj.2004.05.012.
 42. Gorg, O. Drews, C. Luck, F. Weiland, W. Weiss, 2-DE with IPGs. *Electrophoresis* 30 Suppl 1 (2009) S122-132. DOI: 10.1002/elps.200900051.
 43. G. Cicek, T. Vuorinen, I. Stahle, P. Stepanek, N. Freudenberg, R. Brandsch, Coxsackievirus B3 infection induces anti-flavoprotein antibodies in mice. *Clin. Exp. Immunol.* 122 (2000) 404-409. DOI: cei1389.
 44. E. Hammer, T. Q. Phong, L. Steil, K. Klingel, M. G. Salazar, J. Bernhardt, R. Kandolf, H. K. Kroemer, S. B. Felix, U. Volker,

- Viral myocarditis induced by Coxsackievirus B3 in A.BY/SnJ mice: analysis of changes in the myocardial proteome. *Proteomics* 10 (2010) 1802-1818. DOI: 10.1002/pmic.200900734.
45. L. A. Taylor, C. M. Carthy, D. Yang, K. Saad, D. Wong, G. Schreiner, L. W. Stanton, B. M. McManus, Host gene regulation during coxsackievirus B3 infection in mice: assessment by microarrays. *Circ. Res.* 87 (2000) 328-334. DOI: 10.1161/01.RES.87.4.328.
 46. Portig, S. Pankuweit, B. Maisch, Antibodies against stress proteins in sera of patients with dilated cardiomyopathy. *J. Mol. Cell. Cardiol.* 29 (1997) 2245-2251. DOI: 10.1006/jmcc.1997.0463.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.98

Secretome differences between the taxonomically related but clinically differing mycobacterial species *Mycobacterium abscessus* and *M. chelonae*

Jagjit S. Yadav* and Manish Gupta

Microbial Pathogenesis and Toxicogenomics Laboratory, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267-0056, USA

Received: 31 July 2012 Accepted: 21 September 2012 Available Online: 26 September 2012

ABSTRACT

Rapidly growing non-tuberculous mycobacteria (NTM) are significant human pathogens which show high inter-species differences in clinical characteristics (virulence, host immune response) during infection even within a given NTM complex. Understanding the differences between the secreted proteomes of the member species for an NTM complex may reveal the basis of their differential virulence and host pathogenesis potential including host immune reactions. In this study, major secreted proteins of the two taxonomically close but clinically differing member species *M. abscessus* and *M. chelonae* of the *M. chelonae*-*M. abscessus*(MCA) complex were compared using an approach based on 2-dimensional gel electrophoresis (2-DE) and MALDI-TOF analyses. The two secretomes showed dramatic differences. Of the 73 major secreted proteins identified, majority were expressed in a species-specific manner, including 37 in *M. chelonae* and 32 in *M. abscessus*. Interestingly, 9 of these differentially expressed proteins were orphan proteins showing homology to either hypothetical proteins or those with no defined function. The other 60 distinctly expressed proteins were homologs of those associated with various bacterial cellular functions and virulence, namely cell wall synthesis or lipid metabolism, metabolic and respiratory pathways, stress response and signal transduction, gene regulation, and immune response. This information on species-specific secreted proteins would help understand the critical virulence factors and host pathogenesis mechanisms in these mycobacterial species and provide the basis for developing better therapeutic strategies. These proteins may also serve as potential targets for species-specific diagnosis as an additional outcome. To our knowledge, this is the first attempt to characterize the secretome of *M. chelonae* (for which the genome sequence is not yet available) and the secretome differences between *M. abscessus* and *M. chelonae*.

Keywords: Secretome; Secreted proteins; 2-DE; *Mycobacterium*; *M. chelonae*; *M. abscessus*.

Abbreviations:

2DE, Two-dimensional gel electrophoresis; RGM, rapidly growing mycobacteria; NTM, non-tuberculous mycobacteria; MCA, *M. chelonae*-*M. abscessus*; SGM, slow growing mycobacteria; MB7H10, Middle brook 7H10; ATCC, American Type Culture Collection.

1. Introduction

Non-tuberculous mycobacteria (NTM) are ubiquitously distributed in hospital and other environments such as drinking water, soil, biofilms, aerosols, swimming pools, and metal working fluids [1]. There have been increasing reports

of NTM infections in both immunocompetent and immunocompromised individuals particularly HIV patients [2, 3]. While majority of the past studies on virulence factors and mechanisms in NTM have been focused on slowly growing

*Corresponding author: Jagjit S. Yadav, Environmental Genetics and Molecular Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, 3223 Eden Avenue, Kettering Building, Cincinnati, Ohio 45267-0056, USA. Phone: number: +1 5135584806; Fax number: +1-5135584397; E-mail address: Jagjit.Yadav@uc.edu

mycobacteria (SGM) group particularly *Mycobacterium avium* complex [3], little is known on the rapidly growing mycobacteria (RGM) species. Among the RGM, the *M. chelonae*-*M. abscessus* (MCA) complex is particularly significant considering its frequency of occurrence in clinical conditions and extraordinary drug resistance. *M. abscessus* and *M. chelonae* are the original members of the MCA complex that are most frequently detected in clinical conditions associated with this complex [3, 4]. Recently, *M. immunogenum* and two former subspecies of *M. abscessus* viz. *M. bolletii*, and *M. massiliense* have been added as new member species under this complex [5, 6, 7]. The MCA species are genetically closely related and thus require development of specialized diagnostic strategies to differentiate their clinical and environmental isolates, as reported in our recent efforts [8, 9].

Species of the MCA complex are opportunistic pathogens that have been reported to cause several pseudo-outbreaks [10] and chronic infections in lung or other tissues (skin, soft tissues, liver, joints, and female genital tract) of the human body [2]. Many of the infections occur inside the body after normal surgery procedures [11-17]. Pulmonary diseases caused by the MCA species [3], manifest as infection and/or immunological pathologies. The infection pathologies are characterized by tuberculosis (TB)-like symptoms [2]. On the other hand, species of this complex have been implicated as etiological agents of the immune-mediated lung disease hypersensitivity pneumonitis (HP) in machinists exposed to contaminated machining fluids [18]. In this context, our laboratory has reported the occurrence of multiple genotypes of MCA species *M. chelonae*, *M. immunogenum* and *M. abscessus*, from these fluids [19, 20].

Despite being taxonomically closely-related, member species of the MCA complex vary in their clinical presentation both in terms of disease potential/severity and drug resistance. *M. abscessus*, in particular, is a newly emerged pulmonary pathogen increasingly being isolated from cystic fibrosis and other respiratory patient populations [21]. It is now the most frequently encountered species in RGM case reports and is the causative pathogen in an estimated 80% cases of the RGM pulmonary disease [22]. Collectively, these clinical epidemiological studies imply a relatively greater potential of *M. abscessus* in causing MCA pulmonary infections. Independent experimental studies on its pathogenesis potential have shown that *M. abscessus* can replicate in vitro in cultured macrophages and multiply or persist in a mouse model of infection [23]. In contrast, *M. chelonae* showed early clearance both in the infected macrophages [23, 24] and a mouse pulmonary infection model (Unpublished data). Other studies using mouse intravenous infection models demonstrated differential immune control (clearance) of *M. abscessus* versus *M. chelonae* infection [25]. Collectively, the above discussed clinical/phenotypic differences between *M. abscessus* and *M. chelonae* are expected to result from the differences in the underlying molecular factors in these species. However, such virulence and host response-determining factors in these species of the MCA complex

have not yet been characterized.

Secreted proteins are known to play important role in the survival of tuberculous species inside the host or in the environment. Several secreted proteins with specific role in virulence have been reported from *M. tuberculosis* and related SGM species [26, 27]. However, little information is available on the secretomes of non-tuberculous mycobacteria particularly the RGM species of clinical significance [22]. We recently published the immunoproteome (including secreted antigens) of the MCA species *M. immunogenum* [28]. The aim of the current study was to identify and compare secreted proteins of the two prominent and clinically significant member species *M. abscessus* and *M. chelonae* of this complex. The experimental approach included proteomic profiling of their culture filtrate proteins based on 2-DE and MALDI-TOF analyses, and comparing the two secretomes for shared and unique proteins and their putative functional relevance. Recently, the complete genome sequence of *M. abscessus* has been reported [29] (<http://www.ncbi.nlm.nih.gov>) whereas the genome sequence of *M. chelonae* is not yet available. This study on secretome analysis in these two closely-related MCA species will complement the post-genomic analysis and help in studying their systems biology for understanding the virulence and host pathogenesis differences. The identified secreted proteins may have importance in the development of novel therapeutics and vaccine candidates.

2. Materials and Methods

2.1 Strains and culture conditions

M. abscessus ATCC 23006, a lung isolate from a patient's sputum, and *M. chelonae* ATCC 35752, the type strain originally isolated from tortoise, were obtained from the American Type Culture Collection (ATCC). Both species were maintained on Middle brook 7H10 (MB7H10) agar and cultivated using Sauton's medium. For proteomic studies, liquid cultures of the two species were grown to log phase under identical conditions, using Sauton's broth and incubation by shaking (250 rpm) at 37 °C.

2.2 Preparation of secretory proteins extract

Bacterial cells from either species were pelleted by centrifugation (3000g) at 4 °C for 30 min and the culture supernatant was clarified by passing through a 0.22 µm membrane filter. The filtrate was concentrated by trichloroacetic acid (TCA) protein precipitation method [28] followed by resuspension of the precipitate in rehydration buffer (9 M urea, 2M thiourea, 4% CHAPS, 65mM DTT) to isolate the secretory protein extract. Protein concentration in the resulting protein extract was determined using Quick Start Bradford protein assay kit (Bio-Rad, Hercules, CA), according to the manufacturer's specifications [30].

2.3 Isoelectric focusing and two-dimensional gel electrophoresis (2-DE)

For the isoelectric focusing (IEF) prestep, 100 µg of the protein extract was added to 125 µl of rehydration buffer (9 M urea, 2M thiourea, 4% CHAPS, 65mM DTT, 0.5% IPG buffer and 0.002% bromophenol blue). The IEF separation was achieved using a 7cm IPG strip pH range 4-7 (GE Healthcare, Piscataway, NJ) on the IPGphore II isoelectric focusing system (GE Healthcare, Piscataway, NJ). The strip was rehydrated using 50µA for 12h at 15°C and then focused for 16000 volt hours using an upper current limit of 50 µA. Prior to running the 2D gel, the IPG strip was incubated twice for 10 min each in equilibration buffer (6 M Urea, 30% glycerol, 0.05 M Tris pH 8.8, 2% SDS, 0.002% bromo phenol blue), first in conjunction with 135 mM DTT and then with 135 mM iodoacetamide.

The IEF separated proteins were subjected to 2-DE using 15% SDS-PAGE gel. The protein gels were stained with SYPRO Ruby (Invitrogen, Carlsbad, CA) and the gel images were taken and processed using the 2D imaging software Imagemaster™ 2D Elitegive version 4.01 (GE Healthcare, Piscataway, NJ). The isolated protein spots were manually picked from the SYPRO Ruby-stained gel with one touch spot picker (The Gel Co., San Francisco, CA) and prepared for mass spectroscopy analysis, as described below.

2.4 MALDI-TOF peptide mass mapping

The well separated protein spots on the 2D gel were excised and chopped into small pieces followed by 3 alternate cycles of washing using 500 µl of 50% acetonitrile (ACN) in 25 mM ammonium bicarbonate buffer pH 8.0 and dehydration using 100% ACN. The treated gel pieces were dried completely by vacuum centrifugation followed by rehydration at 4°C for 1h with 25 µl trypsin digestion buffer (100 mM ammonium bicarbonate buffer and 1 mM CaCl₂) containing 50 ng trypsin gold (Promega, Madison, WI, USA). Extra trypsin buffer was then added for complete submerging of the gel and the reaction mixture incubated at 37°C with continuous shaking for 16h. Following centrifugation, the individual gel slices were subjected to protein elution for 10 min using 50µl of 5% trifluoroacetic acid (TFA) in 50% ACN (v/v). This process was repeated and the extracts were pooled. The eluted digested protein sample was then evaporated by vacuum centrifugation and kept at -20°C until use. Just before analysis, the digested protein samples were dissolved using 5% TFA in 50% ACN.

The MALDI-TOF analysis was performed on a PE Voyager DE_STR biospectrometry work station (Applied Biosystems, Foster city, CA) set in reflector mode in the mass range 500 to 5000 Da. The matrix α-cyano-4-hydroxycinnamic acid (CHCA) was prepared in a saturated solution of 50% ACN-0.1% TFA. For the mass spectral analysis, equal volumes of the matrix solution (1µl) and the sample (1µl) were mixed on the sample plate, and air dried to form crystals.

Each peptide mass spectrum was calibrated using external standards from Sigma (Insulin oxide β, ACTH Fragment, Angiotensin I, Bradykinin) and internally calibrated with trypsin autolysis peaks; mono isotopic peaks of trypsin auto digests were 842.508, 1045.504 and 2211.108, 2225.12 m/z.

2.5 Bioinformatic analysis

An automated analysis of the mass peaks was done against the protein database using MASCOT search engine (http://www.matrixscience.com/search_form_select.html). To assign a positive identification, at least three peptides had to match, with a search tolerance of 100 ppm while allowing one miscleavage as used in a previous study [28]; possible fixed modifications ascribed to alkylation of cysteine by carbamidomethylation and oxidation of methionine were taken into consideration. For the probability based peptide mass fingerprinting identification, a minimum of one significance hit (P<0.05) was considered as an identity. In some instances, the protein was identifiable despite the low mass score because its top hit was a mycobacterial protein (which allowed it to be differentiated from other proteins). The bioinformatics approach for protein identification is demonstrated in the online Supplementary Information showing tRNA pseudouridine synthase protein identification of *M. chelonae* as an example (supplementary file 1).

2.6 Mycobacterial database searching

The MASCOT-identified proteins were characterized using the available mycobacterial databases namely Tuberculist (www.sanger.ac.uk), Proteome 2D-PAGE Database (<http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page>), the recently released whole genome sequence database for *M. abscessus* (EMBL accession numbers: CU458896, chromosome; CU458745, plasmid), and the published information on mycobacterial and other bacterial proteins. Function-based profiling of the identified secreted proteins was done on the basis of available database and literature mining, mainly using the functional annotation database developed by the Sanger Centre of Genomic Research (http://www.sanger.ac.uk/Projects/M_tuberculosis/Gene_list) and the published literature retrieved via PubMed.

3. Results and Discussion

Investigation of secretomes offers a unique tool to investigate the pathogen factors playing a direct role in host-pathogen interaction and clinical manifestation of infection. This study compared the secretome profiles of *M. abscessus* and *M. chelonae* based on 2-DE and MALDI-TOF analysis of the culture filtrates. The comparisons were made in terms of differential secretion of the protein factors known to be of relevance in pathogen virulence and host response including immune reaction in mycobacteria and other bacterial pathogens (see Tables 1 and 2 and Figures 1-3).

3.1 Optimization of 2-DE separation of the secreted proteins

For efficient comparative secretome analysis, the 2-DE separation of the culture filtrate proteins from the two mycobacterial species was first optimized by varying the IEF pH range and SDS-PAGE concentrations. We observed a better resolution of the proteins in the acidic pH range 4-7 than in a broader pH range (pH 3-10) and in 15% SDS-PAGE, for both the mycobacterial species. More than 100 proteins were visualized on the SYPRO Ruby-stained gel, run in triplicate,

for either species (Figures 1A and 1B). Major protein spots manually excised from the 2D gels and trypsin digested were reliably identified by MALDI-TOF mass spectroscopy and database searching (Tables 1 and 2).

3.2 Functional grouping of the secreted proteins

Majority of the identified secreted proteins were found to be the homologs of the proteins that play diverse roles in both pathogenic and nonpathogenic bacteria. These identi-

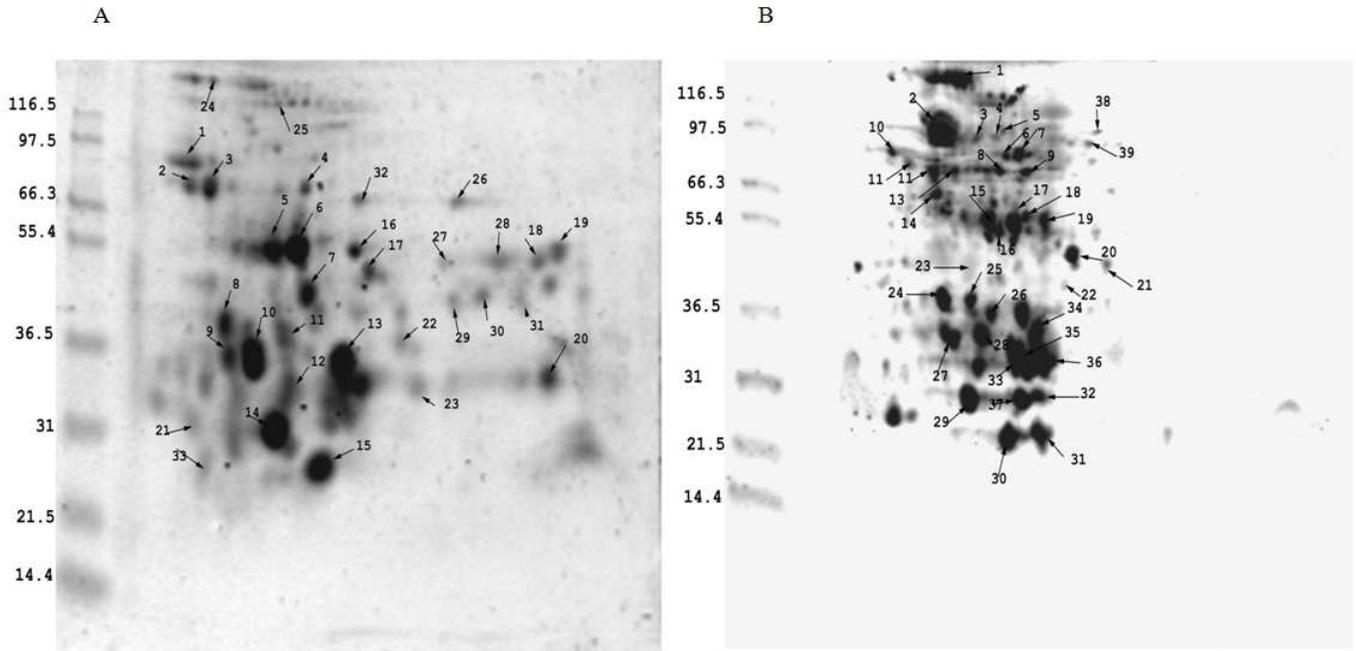


Figure 1. Two-dimensional gel electrophoretic (2-DE) separation of the secretome proteins (A). *M. abscessus* (ATCC 23006). (B). *M. chelonae* (ATCC 35752). Culture filtrate proteins prepared from the actively growing log-phase cells of either species were separated on 2-DE gels (IEF pH gradient 4-7) and stained with SYPRO Ruby as described under Materials and Methods. The approximate positions of molecular weight markers (kDa) are indicated. The spot numbers indicated on the gel correspond to the protein numbers presented in Tables 1 and 2.

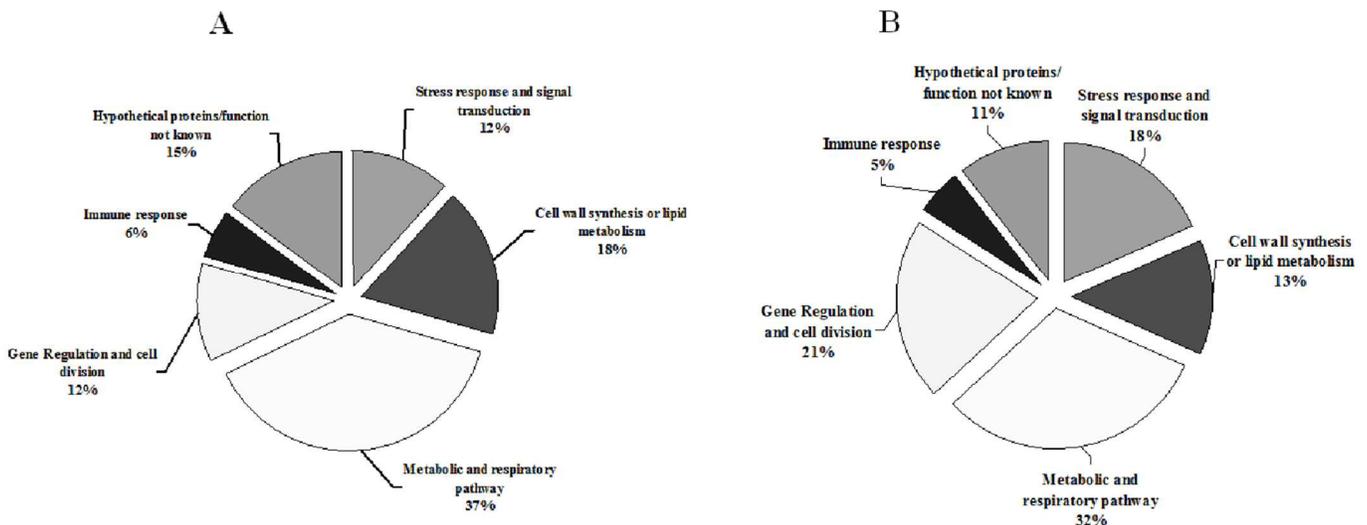


Figure 2. Functional distribution of the identified secreted proteins. (A) *M. abscessus* (ATCC 23006). (B) *M. chelonae* (ATCC 35752).

fied protein homologs in the two species could be categorized in six different functional groups (Figures 2A and 2B), namely (1). Cell wall synthesis or lipid metabolism; (2) Stress response and signal transduction; (3). Metabolic and respiratory pathways; (4). Gene regulation and cell division; (5). Immune response ;(6). Hypothetical proteins of unknown function. Some of the secreted proteins could be assigned to more than one of these functional categories. Relative percent distribution of the individual functional categories in the two secretomes is presented in Figure 2.

3.3. Comparison of the *M. abscessus* and *M. chelonae* secretomes

The 2-DE profiles on the culture filtrate showed that *M. abscessus* and *M. chelonae* secrete about the same range of detectable proteins in the extracellular culture environment (Figure 1). However, a comparison of the two profiles led to an identification of species-specific (unique) and common (shared) proteins (Figure 3). A total of 39 secreted proteins were identified in *M. chelonae*, of which 37 were unique to this species secretome whereas 2 represented shared proteins

between the two species. In *M. abscessus*, 34 major proteins were identified, of which 32 were unique to this species secretome.

3.3.1. Shared secreted proteins

The two secreted proteins common between the *M. abscessus* and *M. chelonae* secretomes viz. elongation factor Tu (EF-Tu) and Acyl-CoA dehydrogenase (FadE6) are known to play roles in survival or virulence in other *Mycobacterium* species. For instance, EF-Tu is a GTPase which helps in binding of aminoacyl-tRNA to ribosomes during protein biosynthesis and shows up-regulation during hypoxia and high iron conditions *in vitro* [31-33]. This protein is also upregulated during mycobacterial infection of macrophages [33]. It has been shown to be associated with the membrane in *M. leprae* [34] and in *E. coli* during starvation indicating that EF-Tu may have role in the regulation of cell growth and the organism's response to stress such as in response to antimycobacterial therapy [35]. FadE6 belongs to the class of flavoproteins which play important role in the oxidation of fatty acyl-CoA [36]. It uses n-octanoyl-CoA as a substrate

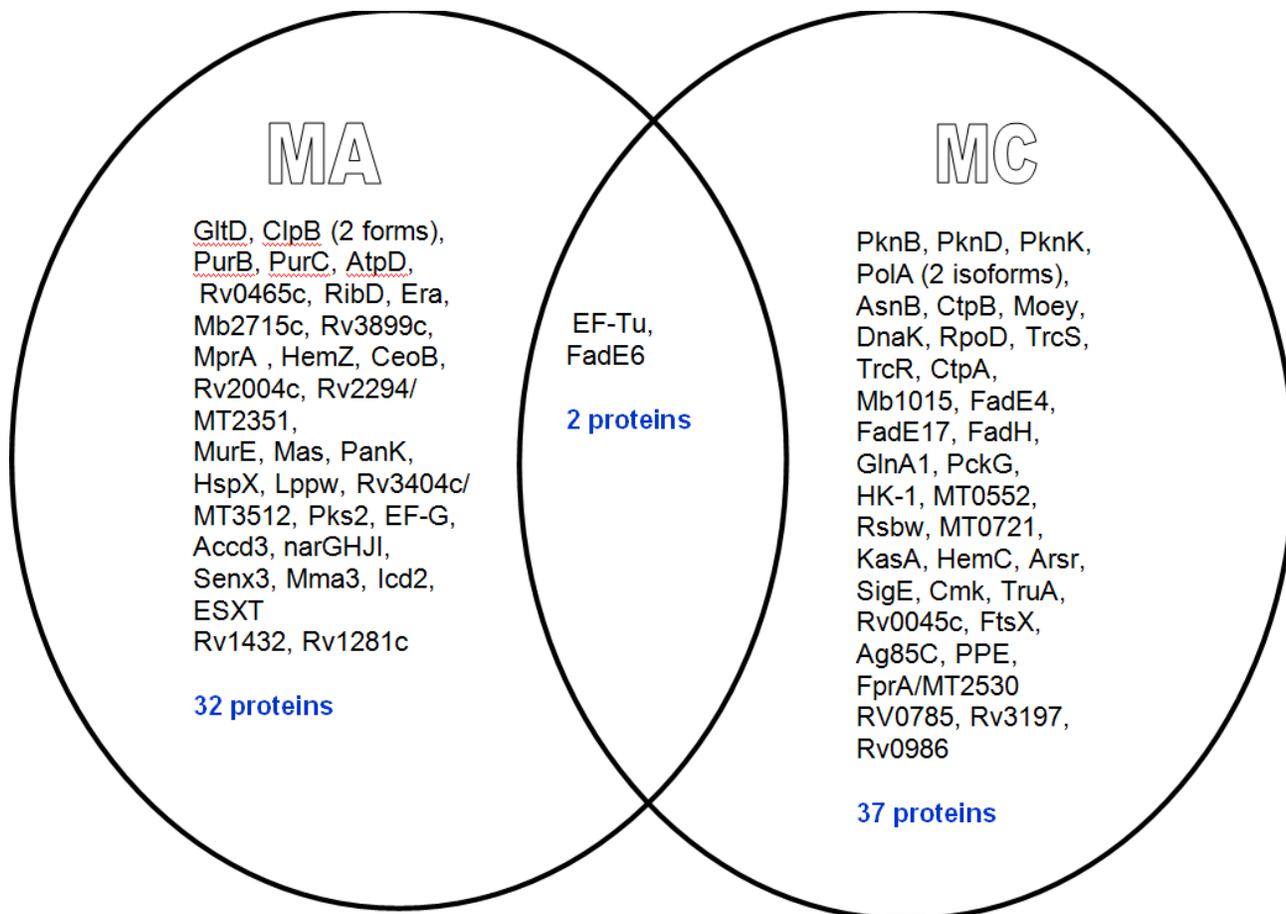


Figure 3. Differential distribution of the secreted proteins in *M. abscessus* versus *M. chelonae*. The two circles in the Venn diagram show the species-specific (unique) proteins for the individual species whereas the overlap region between the two circles shows the proteins common in both species.

Table 1. Identified secretory proteins of *M. abscessus*

Spot no	Protein name	NCBI accession number	MW /pI	Sequence coverage (%)	Number of mass values matched	MOWSE score*	Putative function	Ortholog identified in <i>M. chelonae</i> ?
1	Glutamate synthase (GltD)	Q7TVH8	53399/5.52	29	9	32	Metabolism of nitrogen	No
2	Chaperone protein ClpB	Q73T66	92632/5.2	30.3	14	40	Preventing protein aggregation and assisting in the refolding of denatured proteins	No
3	Chaperone protein ClpB	Q73T66	92632/5.2	18	10	25	Involved in protein folding	No
4	Probable acyl-CoA dehydrogenase (FadE6)	B70628	78428/5.27	21	10	25	Acyl-CoA dehydrogenase activity (Fatty acid metabolism)	Yes
5	Adenylosuccinate lyase (PurB)	P71832	51004/5.94	45	23	28	Involved in purine biosynthesis pathway	No
6	ATP synthase subunit beta (AtpD)	B70775	53061/4.86	43	8	27	ATP generation	No
7	Probable transcription regulator (Rv0465c)	CAA17420	53039/6.63	45	23	28	Gene regulation	No
8	Riboflavin biosynthesis (RibD)	P71677	35345/6.40		6	28	Riboflavin biosynthesis	No
9	Bex GTP-binding protein (Era homolog)	P0A562	32440/ 5.76	41	10	33	Essential for cell growth	No
10	Conserved hypothetical alanine, glycine and valine rich protein (Mb2715c)	Q7TY31	27199/5.35	36	10	20	Unknown function	No
11	Hypothetical protein (Rv3899c)	D70599	40763/5.67	28	6	28	Unknown function	No
12	Response regulator protein MprA	Q84BX0	25908/5.29	15	5	27	Signal transduction	No
13	Probable Heme protein (HemZ)	H70710	37121/5.69	27	5	27	Heme biosynthesis	No

Table 1. Identified secretory proteins of *M. abscessus* (Continuation)

14	Probable PurC protein	G70708	32910/5.12	34	10	26	Purine biosynthesis	No
15	K ⁺ transporter protein (CeoB)	Q7D6R7	24195/5.97	52	7	25	K transport/ drug resistant/ survival in macrophage	No
16	Hypothetical protein (Rv2004c)	B70759	54389/5.89	36	12	21	Unknown function	No
17	Aminotransferase (At) (Rv2294/MT2351)	P63502	44256/5.78	38	7	37	Possible role in nitrogen metabolism	No
18	UDP-N-acetylmuramoylalanyl-D-glutamate --2,6-diaminopimelate ligase (MurE)	P65478	55307/ 6.29	29	12	22	Biosynthesis of cell wall	No
19	Probable dehydrogenase (Rv1432)	Q7D8G3	50445/9.24	23	11	20	Function not known	Yes
20	Putative mycocerosyl (Mas)	Q02277	29808/11.39	66	17	33	Synthesis of branched chain fatty acids	No
21	Type III pantothenate Kinase (PanK)	Q7TW42	29287/5.99	36	12	23	Coenzyme A biosynthesis	No
22	14 kDa antigen HspX	P0A5B8	16217/ 5.00	68	11	28	Antigenic protein	No
23	Lipoprotein precursor (Lppw)	P65305	33381/5.18	23	6	18	An alanine-rich protein found in cell membrane	No
24	Protein Rv3404c (MT3512)	P65073	26498/5.55	52	11	33	Function not known	No
25	Probable polyketide synthase 2 (Pks2)	Q7D4T0	225636/5.29	13	29	21	Lipid metabolism	No
26	Elongation factor G (EF-G)	P0A556	77154/4.96	30	12	32	Protein synthesis	No

Table 1. Identified secretory proteins of *M. abscessus* (Continuation)

27	ABC transporter ATP-binding protein (Rv1281c/MTI318)	P63395	65253/6.02	21	10	27	Binds with ATP and helps in transport of oligopeptide	Yes
28	Acetyl-CoA carboxylase Accd3 protein (Accd3)	E70783	51740/6.10	15	10	21	Mycolic acid biosynthesis	No
29	Elongation factor Tu	P0A559	43566/5.28	35	8	21	Protein synthesis	Yes
30	Nitrate reductase (narGHII)	NC_002755	62073/6.02	25	11	28	Nitrate assimilation	No
31	Sensor-like histidine kinase (Senx3)	P0A601	44797/6.06	19	4	18	Signal transduction	Yes
32	Methoxy mycolic acid synthase 3 (Mma3)	Q8VKHI	34034/5.35	55	16	30	Mycolic acid biosynthesis	No
33	Probable isocitrate dehydrogenase protein (Icd2)	C70848	82498/5.24	18	8	24	Involved in TCA cycle.	No
34	Esat-6 like protein ESXT	O06261	11114/5.89	60	4	19	Immunogenic protein	No

Table 2. Identified secretory proteins of *M. chelonae*

Spot no	Protein name	NCBI accession number	MW /pI	Sequence coverage (%)	Number of mass values matched	MOWSE score*	Putative function	Ortholog identified in <i>M. abscessus</i> ?
1	Probable serine/threonine-protein kinase (PknK)	Q7TXA9	119373/5.52	8	6	15	Signal transduction	No
2	DNA polymerase I (Pol A)	P46835	99793/5.1	34.5	15	87522	DNA replication	No
3	DNA polymerase I (Pol A)	P0A551	98473/5.0	27.0	9	1.01e+7	DNA replication	No

Table 2. Identified secretory proteins of *M. chelonae* (Continuation)

4	Serine/threonine-protein kinase (PknB)	D70699	66469/5.22	19	6	28		Signal transduction	No				
5	Probable asparagine synthetase (AsnB)	B70785	72104/6.30	36	15	40		Asparagine biosynthesis	No				
6	Cation-transporting P-type ATPase B (CtpB)	P46840	72916/4.79	21	6	26		Transport	No				
7	Probable Moey protein kinase	B70741	78133/5.69	12	4	23		Molybdopterin biosynthesis	No				
8	Serine/threonine-protein kinase (PknD)	P95308	69502/5.53	15	4	18		Signal transduction	No				
9	FAD-binding dehydrogenase (Rv0785)	P71838	61258/9.20	30	11	34		Unknown function	Yes				
10	Chaperone protein DnaK	CAA41306	65567/5.16	19	10	28		Heat shock/ chaperonic protein	No				
11	RNA polymerase sigma factor D (RpoD)	P0A603	57765/4.72	21	11	25		Transcription initiation	No				
12	Probable two component sensor protein TrcS	C70624	54557/5.45	19	4	27		Signal transduction	Yes				
13	Probable CtpA protein	D70750	79314/6.69	23	12	36		Copper transport	No				
14	Possible conserved exported protein (Mb1015)	Q7U0W8	42715/5.16	46	6	27		Function unknown	No				
15	Probable FadE17 protein	B70636	44656/5.98	26	10	32		Lipid metabolism	No				
16	Glutamine synthetase (GlnA1)	H70775	53536/5.04	18	4	26		L-glutamine biosynthesis	No				

Table 2. Identified secretory proteins of *M. chelonae* (Continuation)

		P65687	67210/4.92	29	13	40	Carbohydrate metabolism	No
17	Phosphoenolpyruvate carboxykinase [GTP] PckG							
18	Probable acyl CoA dehydrogenase (FadE4)	E70962	63455/5.72	17	7	23	Lipid degradation	No
19	Translation elongation factor Tu (EF-Tu)	A44795	43566/5.28	40	9	30	Protein synthesis	Yes
20	Probable conserved ATP-binding protein ABC transporter (Rv3197)	O53343	49106/6.57	19	6	30	Transporter	Yes
21	Sensor histidine kinase.-HK1	Q7D9B7	52001/6.58	21	5	22	Signal transduction	Yes
22	Hypothetical protein (MT0552)	Q8VKJ9	43068/6.81	24	5	19	Unknown function	No
23	Probable Rsbw protein	H70980	25877/5.39	21	4	26	Acts as an anti-Sigma factor Sig B	No
24	FMN-dependent alpha-hydroxy acid dehydrogenase (MT0721)	Q8VKG1	44442/6.56	15	3	26	Hydroquinone synthesis	No
25	Ketoacyl-ACP synthase A (KasA)	P63455	43289/5.11	9	3	18	Mycolic acid biosynthesis	No

Table 2. Identified secretory proteins of *M. chelonae* (Continuation)

			9	3	18	Mycolic acid biosynthesis	No
25	Ketoacyl-ACP synthase A (KasA)	P63455	43289/5.11	3	18		No
26	Acyl-CoA dehydrogenase (FadE6)	C70825	41425/5.29	5	32	Fatty acid metabolism	Yes
27	Porphobilinogen deaminase (HemC)	A5TZP0	31919/4.84	3	30	Porphyrin biosynthesis	No
28	ABC transporter ATP-binding protein (Rv0986/MT1014)	O53899	27356/5.65	5	22	Role in host interaction	Yes
29	Transcriptional regulator, Arsr family.	Q7DA08	24077/5.13	4	33	Regulation of gene expression	No
30	Sigma factor E (SigE)	P66808	24210/4.98	2	18	Regulation of gene expression	No
31	Cytidylate kinase (Cmk)	Q81FS2	25227/5.06	5	37	Metabolism of nucleotide.	No
32	tRNA pseudouridine synthase A (TruA)	P65847	32600/9.98	7	60	Formation of pseudouridine	Yes

Table 2. Identified secretory proteins of *M. chelonae* (Continuation)

33	Hypothetical protein (Rv0045c)	E70912	32111/5.24	43	8	27	Possible hydrolase	No
34	Cell division protein FtsX homolog	Q7TX91	32771/8.83	19	4	21	Transporter	No
35	Antigen 85-C precursor (Ag85C)	P0A4V5	36748/5.92	13	3	14	Act as a mycolyl transferase	No
36	Probable two-component regulatory protein (TrcR)	D70624	29191/5.82	14	3	13	Signal transduction	No
37	PPE protein	O06246	19799/5.32	24	3	32	Immunogenic protein	No
38	Probable NADPH - dependent 2,4-dienoyl-CoA reductase (FadH)	Q7U0G7	72827/9.62	16	4	26	Lipid biosynthesis	No
39	Ferredoxin oxidoreductase, α subunit (FprA) (MT2530)	Q7D742	69597/5.29	15	5	23	Possible role in iron metabolism or cytochrome P450 activity	No

*Mowse score: The Mowse score, used in peptide mass fingerprinting, is a "similarity score" derived using a statistical model that calculates the "probability of matching N peaks by random chance".

and acts as a medium-chain acyl-CoA dehydrogenase in *M. tuberculosis* [37]. The shared secreted proteins could be the basis of common virulence characteristics in the two species.

3.3.2. Species-specific secreted proteins

It is interesting that *M. abscessus* has the ability to multiply, persist, and cause infection in *in vivo* models [23, 24, 25, Yadav et al unpublished data] whereas *M. chelonae* gets cleared relatively early during infection in these models. It is likely that the unique secreted proteins of *M. abscessus* identified in this study enable these differences in survival and/or virulence of *M. abscessus* inside the host as compared to *M. chelonae*. Specifically, differential secretion of the following major categories of mycobacterial proteins in the two species may be responsible for the species-specific differences in their immunogenicity/virulence and host response characteristics.

3.3.2.1. ESAT-6 secretion system (ESX) proteins.

ESX is a type VII secretion system known to exist in Mycobacteria and many other actinobacteria and Gram-positive bacteria. It comprises of 5 paralogs (ESX-1 through ESX-5) in *M. tuberculosis*. Of these, ESX-1 which is responsible for secretion of its prototypical members Early Secreted Antigen 6 kDa (ESAT-6) and 10 kDa culture filtrate protein (CFP-10), has been shown to play an important role in pathogen virulence and survival inside the host. For instance, the ESX-1 components ESAT-6 (ESXA) and CFP 10 (ESXB) form a complex and allow the TB pathogen to survive within macrophages [38]. In contrast, the vaccine strain *M. bovis* BCG which has ESAT-6 deletion in the RD-1 region [39] is unable to persist in the host. Some of the other four paralogous ESX systems are also known to be essential for pathogen growth. In light of these facts about the functional significance of mycobacterial ESX proteins, it is interesting that we detected a putative ESAT-6 like protein ESXT in *M. abscessus* and not in *M. chelonae* culture filtrate. Besides its relevance in conferring differential virulence potential, the ESAT-6 has proven to be a specific diagnostic target for tuberculosis infection such as in QuantiFERON[®]-TB Gold Test and certain multiplex assays. Likewise, the differentially secreted ESX protein ESXT in *M. abscessus*, could be a potential target for designing species-specific diagnostic assays. In this context, ESXA and ESXB have been previously reported as effective targets for specific diagnosis of *M. abscessus* infection [40]. ELISA-based analysis for these proteins enabled successful diagnosis of infection (abscess) in *M. abscessus*-infected patients (culture positive for *M. abscessus*) but not in the *M. chelonae*-infected patients (culture positive for *M. chelonae*) [40].

3.3.2.2 Cell wall synthesis or lipid metabolism proteins.

Several secreted proteins identified in the two species were

found to be the homologs of proteins/enzymes involved in biosynthesis or degradation of fatty acids (mycolic acids) and glycolipids, the two key components of the mycobacterial cell wall. Except for the fatty acid oxidizing protein FadE6 which is secreted by both *M. abscessus* and *M. chelonae*, others in this functional category were differentially secreted. For instance, *M. abscessus*-specific secreted proteins were Mycocerosyl (Mas), Polyketide synthase 2 (Pks2), Acyl-CoA carboxylase 3 (AccD3), and Methoxy mycolic acid synthase 3 (MmaA3) whereas those differentially detected in *M. chelonae* secretome were Ketoacyl-ACP synthase A (KasA) and NADPH-dependent 2, 4- Dienoyl-CoA reductase (FadH). In addition, the alanine-rich lipoprotein LppW was detected among the major secreted proteins of *M. abscessus* similar to that observed in TB pathogen [26]. Secretion of the predominantly cell envelope-associated lipoproteins has been ascribed to either shedding (release of acylated lipoproteins) or shaving (proteolytic cleavage) in a recent study on *M. tuberculosis* [27].

It is noteworthy that the cell wall biosynthesis gene *mas* (fatty acid synthase) which is considered unique to pathogenic species of the SGM group [36, 41] was expressed in the *M. abscessus* secretome. Expression of this gene is known to upregulate intracellular growth and survival within macrophages [42]. Another cell wall biosynthesis protein Polyketide synthase 2 (Pks2) is known to be involved in the formation of methyl-branched fatty acyl components of sulfolipids [43] such as sulfolipid-1 (SL-1) and plays a role in establishment of TB infection in human host [44]. Homologs of the proteins involved in biosynthesis of mycolic acids (that form the hydrophobic outer layer of mycobacterial cell wall) showed differential distribution in *M. abscessus* (Acyl-CoA carboxylases 3 (AccD3), Methoxy mycolic acid synthase 3 (MmaA3)) versus *M. chelonae* (Ketoacyl-ACP synthase A (KasA)) secretomes. The FAS pathway protein NADPH-dependent 2, 4- Dienoyl-CoA reductase (FadH) differentially secreted in *M. chelonae* is known to act as a link between the FAS-I and FAS-II pathways [45].

3.3.2.3. Stress response proteins

In response to stresses such as heat, oxidizing conditions, and toxicant exposures, the cells produce a set of proteins loosely referred to as heat shock proteins (HSPs). Some of these proteins act as molecular chaperones which bind with and stabilize proteins at the intermediate stages, including folding, assembly, translocation across membranes, and degradation. In our analysis we found two differentially secreted chaperonic proteins ClpB (*M. abscessus*) and DnaK (*M. chelonae*). ClpB is a heat shock protein [46] of the clp protein family that acts as an ATP-dependent protease. Notably ClpB separated as two distinct spots on the 2-DE gel (Figure 1; Table 1). Since the two spots corresponded to the same protein accession number, these may be the length variants of the same protein. DnaK is a major immunodominant antigen in pathogenic mycobacteria [47, 48] and is found to be

upregulated when cells are subjected to high temperature. It is considered as an inhibitor of heat shock response to enable survival of the cells. Other stress proteins detected in *M. chelonae* but not in *M. abscessus* included an RNA polymerase sigma factor E (SigE), three serine-threonine kinases STPKs (PknB, PknD, and PknK), and the TrcRS two-component system proteins TrcR and TrcS. Additional histidine kinase protein homologs differentially detected in the two species were HK1 (*M. chelonae*) and Senx3 (*M. abscessus*). Differential secretion of the chaperonic and other heat shock proteins in the two species implies their variable potential to survive and respond under physiological stress conditions, which in turn may contribute to their phenotypic diversity in terms of host-pathogen interactions.

3.3.2.4. Metabolic and respiratory pathways associated proteins.

This functional category encompassed the largest fraction (32-37%) of the identified secretory proteins in both species (Figure 2). Differential distribution of the proteins of this category in the two species was as follows: *M. abscessus*-Adenylosuccinate lyase (PurB) and Phosphoribosylamidoimidazole- succino-carboxamide synthase (PurC), F1Fo ATP synthase subunit beta (AtpD), Riboflavin biosynthesis enzyme (RibD), Nitrate reductase (narGHJI), Isocitrate dehydrogenase 2 (Icd2), Phosphoenol pyruvate carboxykinase (PckG), TrkA protein (also designated as CeoB), Aminotransferase (At); *M. chelonae*- Asparagine synthase B (AsnB), Copper-transporting ATPase (CtpA), FMN-dependent alpha-hydroxy acid dehydrogenase (MmcS), Porphobilinogen deaminase (PBGD), Pseudouridine synthase (TruA), Pseudouridine synthase (TruA).

ABC transporter class of proteins important in various cellular processes with likely role in drug resistance, immunity, and/or pathogenesis [49] were detected in both the secretomes albeit with a differential distribution of the members in *M. abscessus* (Rv1281c) versus *M. chelonae* (Rv0986, Rv3197). Such distribution between the two species implies their role in differentially conferring multiple virulence characteristics to these pathogens.

Notably, the two key nitrogen metabolism pathway enzymes Glutamate synthetase (GltD) and Glutamine synthetase (GlnA1) were differentially detected in *M. abscessus* (GltD) and *M. chelonae* (GlnA1) secretomes. GltD is known to be involved in glutamate synthesis whereas GlnA1 is responsible for the incorporation of ammonia into glutamate to make glutamine at low ammonia concentration [50]. GlnA1 is a major secreted protein in the culture filtrate of *M. tuberculosis* (representing approximately one-third of its total measurable enzyme activity), a feature that is considered highly specific in pathogenic mycobacteria. It is one of the important drug targets in pathogenic mycobacteria considering its crucial roles in pathogen survival inside the phagosome [51]; this occurs via modulation of ammonia levels, which may in turn influence phagosomal pH and phago-

some-lysosome fusion and cell wall formation [52]. L-glutamine is considered a major component of the cell wall of pathogenic but not nonpathogenic mycobacteria. Its presence in the *M. chelonae* secretome may be one of the contributing factors to its drug resistance.

3.3.2.5. Gene regulation and cell division related proteins

The member proteins of this functional category were differentially distributed between the two species as follows: *M. abscessus* (Elongation Factor G, GTP-binding protein Era homolog); *M. chelonae* (DNA polymerase I or PolA, RNA polymerase sigma factor D (RpoD), Anti-sigma factor (RsbW), Transcription regulatory protein (ArsR), the cell division protein FtsX). Notably PolA separated as two distinct isoforms with distinct accession numbers (Table 2). Being associated with basic cellular functions, these proteins could serve as important drug target candidates.

3.3.2.6 Immune response-related proteins:

An estimated 5-6% of the secreted proteins in the two secretomes corresponded to those related to immune response (Figure 3). Differential secretion of these proteins was as follows: *M. abscessus* (ESAT-6, 14 kDa antigen HspX), *M. chelonae* (Antigenic protein 85C (Ag85C), PPE protein).

In addition to the ESAT-6 like protein shared between the two secretomes, *M. abscessus* differentially secreted HspX14 (also known as 16 kDa antigen or HSP 16.3). This protein is required for the growth of mycobacterium in macrophages and is upregulated under anaerobic conditions [53]. It has been proposed to play a negative regulatory role in the multiplication of the TB pathogen during *in vivo* infection [54]. It induces humoral immune response in TB patients and also induces T cell and B cell immune responses in latent infections [55, 56]. In contrast, *M. chelonae* secretome consisted of the major mycobacterial antigens Ag85C and PPE. Ag85C possesses a mycolyltransferase enzyme activity and plays an important role in cell wall biosynthesis and survival in mycobacteria [57]. It forms complex with Ag85A and Ag85B and constitutes a major cell wall component. Secreted protein Ag85B has been used for the diagnosis of tuberculosis [58]. It induces strong Th1 response and protects mice from the infection when used for immunization along with Ag85A or with other known mycobacterial antigens [59]. The other secreted protein under this functional category was a homolog of PPE59 (O06246) which belongs to the large PPE multigene family of antigens in mycobacteria. This family constitutes around 10% of the *M. tuberculosis* genome [60]. Members of this protein family are rich in asparagine and glycine and contain multiple repeats of AsnXGlyXGlyXAsnXGly signature as a major polymorphic tandem repeat (MPTR) [60]. Consequently, these proteins possess fewer tryptic cleavage sites making them difficult to detect in global proteomic analysis. Our observation is significant considering that

there are fewer chances of detecting them in culture filtrate as compared to the cell wall fraction. For instance, in a latest otherwise exhaustive global proteomic study on the TB bacillus, only one PPE protein could be detected in the culture filtrate (27). Because of the presence of sequence variation at the C-terminal, it has been suggested that the PPE family proteins are responsible for the antigenic variation in mycobacteria [61]. Differential expression of an extracellular PPE may therefore be responsible for conferring yet unknown differential immunogenic characteristic(s) to *M. abscessus*.

In our recent immunoproteomic study on another species of the MCA complex *M. immunogenum* [28], we detected 4 secreted antigens, including EF-Tu (similar to *M. chelonae* and *M. abscessus*), antigen 85A and CtpA (similar to *M. chelonae*) and OtsB (unlike either species in the current study).

3.4. Proteins with no known function (orphan proteins)

In this study, 9 proteins with yet unknown function (orphan proteins) were identified from the secretomes of *M. abscessus* (15% of the total) and *M. chelonae* (11% of the total) (Figure 2). Seven of the orphan proteins are homologs of the *M. tuberculosis* hypothetical proteins namely Rv3899c, Rv2004c, Rv3404c, Mb2715c, Mb1015, MT0552, and Rv0045c (Tables 1 and 2). Two of the orphan proteins viz. FAD-binding dehydrogenase (Rv0785) in *M. chelonae* (Table 2) and a probable dehydrogenase (Rv1432) in *M. abscessus* (Table 1), matched properly named proteins in the database.

4. Concluding remarks

In conclusion, this study provides the first account on the major secretory proteins in the RGM species *M. chelonae* for which the genome sequence is not yet known. Importantly, the study revealed dramatic differences between the secretomes of the two taxonomically close RGM species, *M. abscessus* and *M. chelonae*, which show differences in host immune control and infection response. Majority of the secreted proteins (69 of 73 proteins) showed species-specific distribution implying the significance of the unique sets of secreted factors in conferring the host infection differences to these clinically important mycobacterial species. Further functional studies on the identified species-specific proteins may elucidate their specific underlying role in conferring the virulence and host response (including host immune reaction) potential in the two RGM species and could provide specific targets for development of anti-RGM therapeutics (drugs and vaccines). As an additional impact of the study, the differentially secreted proteins could also be exploited as targets for species differentiation and development of species-specific infection diagnosis tools.

5. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/96/0>.

Supplementary Material includes Supplementary file 1: Bioinformatic analysis for protein identification based on Peptide mass fingerprinting and Mascot search.

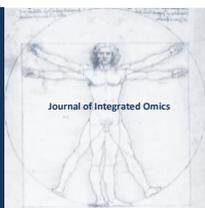
Acknowledgements

The study was supported by the grant 2R01OH007364 (to J.S.Y.) from the National Institute of Occupational Safety and Health, Center for Disease Control and Prevention. The authors acknowledge Drs. George Smulian and Francisco Gomez of the University's V.A Hospital and Medical Center for their guidance in MALDI-TOF analysis and Dr. Ying Wai Lam for help in 2D-gel image analysis.

References

1. T.C. Covert, M.R. Rodgers, A.L. Reyes, G.N. Stelma Jr, Appl. Environ. Microbiol. 65 (1999) 2492-2496.
2. D.E. Griffith, T. Aksamit, B.A. Brown-Elliott, A. Catanzaro, C. Daley, F. Gordin, S.M. Holland, R. Horsburgh, G. Huitt, M.F. Iademarco, M. Iseman, K. Olivier, S. Ruoss, C.F. von Reyn, R.J. Wallace Jr, K. Winthrop, Am. J. Respir. Crit. Care 175 (2007) 367-416.
3. B.A. Brown-Elliott, R.J. Wallace Jr, Clin. Microbiol. Rev. 15 (2002) 716-746.
4. M. Moore, J.B. Frerichs, J. Invest. Dermatol. 20 (1953) 133-169.
5. R.W. Wilson, V.A. Steingrube, E.C. Bottger, B. Springer, B.A. Brown-Elliott, V. Vincent, K.C. Jost Jr, Y. Zhang, M.J. Garcia, S.H. Chiu, G.O. Onyi, H. Rossmoore, D.R. Nash, R.J. Wallace Jr, Int. J. Syst. Evol. Microbiol. 51 (2001) 1751-1764.
6. T. Adékambi, M. Reynaud-Gaubert, G. Greub, M.J. Gevaudan, B. La Scola, D. Raoult, M. Drancourt, J. Clin. Microbiol. 42 (2004) 5493-5501.
7. T. Adékambi, P. Berger, D. Raoult, M. Drancourt, Int. J. Syst. Evol. Microbiol. 56 (2006) 133-143.
8. I.U.H. Khan, S.B. Selvaraju, J.S. Yadav, J. Clin. Microbiol. 43 (2005) 4466-4472.
9. S.B. Selvaraju, I.U.H. Khan, J.S. Yadav, Mol. Cell. Probes 19 (2005) 93-99.
10. D.B. Blossom, K.A. Alelis, D.C. Chang, A.H. Flores, J. Gill, D. Beall, A.M. Peterson, B. Jensen, J. Noble-Wang, M. Williams, M.A. Yakrus, M.J. Arduino, A. Srinivasan, Infect. Control Hosp. Epidemiol. 29 (2008) 57-62.
11. A. Baisi, M. Nosotti, B. Chella, L. Santambrogio, Transpl. Int. 18 (2005) 1117-1119.
12. M. Brickman, A.A. Parsa, F.D. Parsa, Aesthetic. Plast. Surg. 29 (2005) 116-118.
13. I. Stelzmueller, K.M. Dunst, S. Wiesmayr, R. Zangerie, P. Hengster, H. Bonatti, Emerg. Infect. Dis. 11 (2005) 352-354.
14. S.H. Chung, M.I. Roh, M.S. Park, Y.T. Kong, H.K. Lee, E.K. Kim, Ophthalmologica 220 (2006) 277-280.
15. Y.P. Siu, K.T. Leung, M.K. Tong, M.K. Lee, Clin. Nephrol. 63 (2005) 321-324.
16. H. Meyers, B.A. Brown-Elliott, D. Moore, J. Curry, C. Truong, Y. Zhang, R.J. Wallace Jr, Clin. Infect. Dis. 34 (2002) 1500-1507.

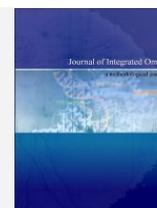
17. V.L. Katz, R. Farmer, J. York, J.D. Wilson, *J Reprod Med* 45 (2000) 581-584.
18. B.G. Shelton, W.D. Flanders, G.K. Morris, *Emerg. Infect. Dis.* 5 (1999) 270-273.
19. I.U.H. Khan, S.B. Selvaraju, J.S. Yadav, *FEMS Microbiol Ecol* 54 (2005) 329-338.
20. R. Kapoor, J.S. Yadav, *FEMS Microbiol. Ecol.* 79 (2012) 392-399.
21. B. Petrini, *APMIS* 114 (2006) 319-328.
22. A-R. Shin, H. Sohn, C.J. Won, B. Lee, W.S. Kim, H.B. Kang, H.J. Kim, S.N. Cho, S.J. Shin, *J. Microbiol.* 48 (2010) 502-511.
23. D. Ordway, M. Henao-Tamayo, E. Smith, C. Shanley, M. Harton, J. Troutd, X. Bai, R.J. Basaraba, I.M. Orme, E.D. Chan, *J. Leukoc. Biol.* 83 (2008) 1502-1511.
24. M.J. Harriff, M. Wu, M.L. Kent, L.E. Bermudez, *Appl. Environ. Microbiol.* 74 (2008) 275-285.
25. M. Rottman, E. Catherinot, P. Hochedez, J-F. Emile, J-L. Casanova, J-L. Gaillard, C. Soudais, *Infect. Immun.* 75(2007) 5898-5907.
26. H. Malen, F.S. Berven, K.E. Fladmark, H.G. Wiker, *Proteomics* 7 (2007) 1702-1718.
27. C. Bell, G.T. Smith, M.J. Sweredoski, S. Hess, *J. Proteome Res.* 11 (2012) 119-130.
28. M.K. Gupta, V. Subramanian, J.S. Yadav, *J. Proteome Res.* 8 (2009) 2319-2330.
29. F. Ripoll, S. Pasek, C. Schenowitz, C. Dossat, V. Barbe, M. Rottman, E. Macheras, B. Heym, J.L. Herrmann, M. Daffé, R. Brosch, J.L. Risler, J.L. Gaillard, *PLoS One* 4 (2009) e5660.
30. M.M. Bradford, *Anal. Biochem.* 72 (1976) 248-254.
31. J. Starck, G. Kallenius, B.I. Marklund, D.I. Andersson, T. Akerlund, *Microbiol.* 150 (2004) 3821-3829.
32. D. Schnappinger, S. Ehrst, M.I. Voskuil, Y. Liu, J.A. Mangan, I.M. Monahan, G. Dolganov, B. Efron, P.D. Butcher, C. Nathan, G.K. Schoolnik, *J. Exp. Med.* 198 (2003) 693-704.
33. L.G. Wayne, H.A. Sramek, *Antimicrob. Agents Chemother.* 38 (1994) 2054-2058.
34. M.A. Marques, S. Chitale, P.J. Brennan, M.C. Pessolani, *Infect. Immun.* 66 (1998) 2625-2631.
35. M. Wilson, J. DeRisi, H.H. Kristensen, P. Imboden, S. Rane, P.O. Brown, G.K. Schoolnik, *Proc. Natl. Acad. Sci. USA* 96 (1999) 12833-12838.
36. D.L. Rainwater, P.E. Kolattukudy, *J. Biol. Chem.* 260 (1985) 616-623.
37. U. Mahadevan, G. Padmanaban, *Biochem. Biophys. Res. Commun.* 244 (1998) 893-897.
38. T. Tan, W.L. Lee, D.C. Alexander, S. Grinstein, J. Liu, *Cell. Microbiol.* 8 (2006) 1417-1429.
39. G.G. Mahairas, P.J. Sabo, M.J. Hickey, D.C. Singh, C.K. Stover, *J. Bacteriol.* 178 (1996) 1274-1282.
40. J.Y. Wang, C.H. Chou, L.N. Lee, H.L. Hsu, I.S. Jan, P.R. Hsueh, P.C. Yang, K.T. Luh, *Emerg. Infect. Dis.* 13 (2007) 553-558.
41. M. Mathur, P.E. Kolattukudy, *J. Biol. Chem.* 267 (1992) 19388-19395.
42. M.S. Li, I.M. Monahan, S.J. Waddell, J.A. Mangan, S. Martin, M.J. Everett, P.D. Butcher, *Microbiol.* 147 (2001) 2293-2305.
43. K. Bhatt, S.S. Gurcha, A. Bhatt, G.S. Besra, W.R. Jacobs, *Microbiol.* 153 (2007) 513-520.
44. P. Kumar, M.W. Schelle, M. Jain, F.L. Lin, C.J. Petzold, M.D. Leavell, J.A. Leary, J.S. Cox, C.R. Bertozzi, *Proc. Natl. Acad. Sci. USA* 104 (2007) 11221-11226.
45. K.H. Choi, L. Kremer, G.S. Besra, C.O. Rock, *J. Biol. Chem.* 275 (2000) 28201-28207.
46. C.L. Squires, S. Pedersen, B.M. Ross, C. Squires, *J. Bacteriol.* 173 (1991) 4254-4262.
47. M.R. Alavi, L.F. Affronti, *J. Leukoc. Biol.* 55 (1994) 633-641.
48. F.M. Collins, J.R. Lamb, D.B. Young, *Infect. Immun.* 56 (1988) 1260-1266.
49. M. Braibant, P. Gilot, J. Content, *FEMS Microbiol. Rev.* 24 (2000) 449-467.
50. G. Harth, D.L. Clemens, M.A. Horwitz, *Proc. Natl. Acad. Sci. USA* 91 (1994) 9342-9346.
51. M.V. Tullius, G. Harth, M.A. Horwitz, *Infect. Immun.* 71 (2003) 3927-3936.
52. G. Harth, P.C. Zamecnik, J.Y. Tang, D. Tabatadze, M.A. Horwitz, *Proc. Natl. Acad. Sci. USA* 97 (2000) 418-423.
53. Y. Yuan, D.D. Crane, R.M. Simpson, Y.Q. Zhu, M.J. Hickey, D.R. Sherman, C.E. Barry 3rd, *Proc. Natl. Acad. Sci. USA.* 95 (1998) 9578-9583.
54. Y. Hu, F. Movahedzadeh, N.G. Stoker, A.R. Coates, *Infect. Immun.* 74 (2006) 861-868.
55. A. Davidow, G.V. Kanaujia, L. Shi, J. Kaviar, X. Guo, N. Sung, G. Kaplan, D. Menzies, M.L. Gennaro, *Infect. Immun.* 73 (2005) 6846-6851.
56. A. Demissie, E.M. Leyten, M. Abebe, L. Wassie, A. Aseffa, G. Abate, H. Fletcher, P. Owiafe, P.C. Hill, R. Brookes, G. Rook, A. Zumla, S.M. Arend, M. Klein, T.H. Ottenhoff, P. Andersen, T.M. Doherty, *Clin. Vaccine Immunol.* 13 (2006) 179-186.
57. J.T. Belisle, V.D. Vissa, T. Sievert, K. Takayama, P.J. Brennan, G.S. Besra, *Science* 276 (1997) 1420-1422.
58. K.S. Senthil Kumar, K.R. Uma Devi, R. Alamelu, *Indian J. Chest Dis. Allied Sci.* 44 (2002) 225-232.
59. S.U. Park, K. Kathaperumal, S. McDonough, B. Akey, J. Huntley, J.P. Bannantine, Y.F. Chang, *Vaccine* 26 (2008) 4329-4337.
60. S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M.A. Quail, M.A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J.E. Sulston, K. Taylor, S. Whitehead, B.G. Barrell, *Nature* 393 (1998) 537-544.
61. G. Delogu, F. Bigi, S.E. Hasnain, A. Cataldi, in: M. Daffe and J.M. Reyat (Eds.), *The Mycobacterial Cell Envelope*, ASM Press, Washington DC, 2008, pp. 133-152.



JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.109

A robust permutation test for quantitative SILAC proteomics experiments

Hien D. Nguyen^{1,2}, Ian A. Wood², Michelle M. Hill^{*1}.

¹The University of Queensland Diamantina Institute, The University of Queensland, QLD 4102 Australia. ²School of Mathematics and Physics, University of Queensland, St. Lucia, QLD 4072 Australia.

Received: 06 September 2012 Accepted: 02 October 2012 Available Online: 14 December 2012

ABSTRACT

Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) along with other relative quantitation methods in proteomics have become important tools in the analysis of cellular and subcellular functions. Although numerous experimental applications of SILAC have been developed, there is no consensus on the use of statistical procedures to analyze the resulting experimental data. SILAC experiments output relative abundance ratios for proteins to quantify differences in cell populations. These ratios have traditionally been analyzed with fold-change methods and hypothesis testing procedures under Gaussian distribution assumptions.

We find that the normality assumption is invalid and can lead to inaccurate quantitation of the significance of differences between cell populations. As a solution, a permutation based hypothesis test as an alternative for assessing significance is introduced. We develop a distribution-free permutation testing methods for assessing various SILAC experiments. These tests generate p-values which can be easily interpreted and if necessary, the false discovery rate of these p-values can be easily controlled. To compare the permutation test against competing methodology, we used a set of simulations based upon a theoretical model of SILAC ratio data.

Through the simulation studies, we find that the permutation test is generally superior to the competing hypothesis tests across the range of simulation scenarios. We also find that the permutation test is typically more powerful and accurate than the competing methods at the five percent level of significance and averaged over the spectrum of significance levels. Because of the broad superiority of the permutation test and the ease of implementation, we propose the use of the permutation test as a standard measure of protein significance in SILAC experiments.

Keywords: SILAC; quantitative proteomics; robust statistics; permutation test; Gaussian mixture model.

1. Introduction

Mass spectrometry-based proteomics methods have become widely used and highly successful tools for the large-scale study of molecular and cellular biology in recent years. Since mass spectrometry analysis for individual samples are performed sequentially, quantitative comparisons are not inherent in these experiments. Over the past decade, several labeling methodologies have been developed to enable quantitative proteomics analysis using multiplex mass spectrometry. The general approach is to label specific amino acids of different samples with different mass ‘tags’, allowing the samples to be mixed for the mass spectrometry analysis, providing relative quantitation in terms of a ratio between the samples. One method popular with cell biologists was

introduced by the laboratory of Matthias Mann in 2002 [1], termed Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC), and involves the incorporation of isotopic amino acids as the cells produce new proteins during their normal growth. Since proteins are labelled during normal cell metabolism, rather than post-preparation of the proteins of interest, as is used in other quantitative proteomics strategies, SILAC offers the potential to account for variations during protein separation, and is particularly suited for subcellular proteomes [2–4] and time-resolved proteomics [2,5]. Although originally developed for cultured mammalian cells, SILAC has now been applied to other species and organisms, including bacteria [6], fly [7,8], plant [9] and yeast [10]. In

*Corresponding author: Michelle M Hill. The University of Queensland Diamantina Institute, Level 4 R-Wing, Princess Alexandra Hospital, Ipswich Road, Woolloongabba QLD 4102 Australia. Telephone: +61 7 3176 7456; Email Address: m.hill2@uq.edu.au

addition, the development of SUPER-SILAC methodology has enabled quantitative analysis with non-metabolic samples, such as archival tissue specimen and plasma [11]. A thorough description of the SILAC method can be found in [12].

In a typical two-plex SILAC experiment, samples are labelled in with 'light' and 'heavy' isotopes by feeding the cell or organism with amino acids containing the required isotopes. Lysates are prepared and then combined based on equal protein content. The target cellular or subcellular proteome are isolated, digested with trypsin, and the peptides are analyzed by tandem mass spectrometry (MS/MS) following one or more separation steps, generally by liquid chromatography (LC). Because of the predictability of the mass shift between the peptides coming from the "light" and "heavy" populations, it is possible to distinguish "heavy" and "light" peptide from the survey mass spectra (MS1), and hence calculate a ratio of their relative intensities. A ratio equal to one would indicate that there are no differences in the abundance of the peptide between the two populations whereas a ratio not equal to one would indicate an up or down-regulation of the peptide.

Furthermore, fragmentation of the parent ion (MS/MS spectra) allows the matching of the quantified peptides to a protein sequence, and the inference of a protein's abundance change by averaging over the matched peptides. Protein inference from MS/MS data is generally performed by specialist software, both commercial and open source. Comprehensive comparisons of these algorithms have recently been performed [13,14]. The current work will focus on the statistical evaluation of such quantitative data.

In a typical SILAC experiment, there may be thousands of protein ratios measured [3,8,12]. Because of the large number of protein ratios quantified, and the existence of errors in the computational and experimental processes, it is highly unlikely that all proteins with ratios different to one are actually differently regulated between the two populations. It is therefore necessary to define a method by which proteins can be considered to be differently regulated or not.

A common approach to assessing protein ratios is through the use of a "fold-change" threshold applied over all protein ratios [1,15]. Under a 1.5 threshold scheme as suggested in [15], a protein is considered differently abundant between the whole populations if the protein ratio is either greater than 1.5 or less than the inverse of 1.5 (2/3). Depending on the experiment, there have been reports of uses of thresholds anywhere between 1.3 [12] to 6 [1].

Because of the differences in variability between experiments, an application of a certain threshold to one experiment may result in a high proportion of false positives whereas the same threshold applied to a different experiment might yield a high proportion of false negatives instead. The combination of variability in results, combined with a need for experimenter specification makes the threshold method less useful for high throughput applications where there is a greater priority to mitigate the proportion of false discover-

ies in an experiment [16,17].

For candidate proteins that will be further validated through orthogonal approaches, such as western blotting, strictly controlling the false discovery rate may not be crucial. However, one of the advantages of omics analyses such as proteomics is the ability to infer pathways and networks from the high throughput data [18,19]. For these workflows, thorough statistical hypothesis testing is crucial in ensuring the true list of altered proteins is carried forward for the network or pathways analyses [20]. There have been a wide variety of hypothesis testing methodologies applied to the assessment of abundance by SILAC ratios such as the one sample t-test [21,22], z-test [23,24], robust z-test (significance A) [17,25,26] and Wilcoxon signed-rank test [22]. The reason for the numerous approaches to hypothesis testing is because of the disagreement about what statistical model and assumption to use for the calculation of p-values. Methods such as the t-test and z-test are based on an assumption that either the protein ratio distribution or peptide ratio is normally distributed.

In this study, we examined the distribution of two-plex SILAC datasets, and demonstrate the assumption violations and deficiencies of the t-test, z-test, robust z-test and Wilcoxon signed-rank test when used with SILAC data. We then propose a new permutation test which does not rely on distributional assumption and demonstrate its utility and superiority over the current methodologies in a simulation study.

2. Material and Methods

Peptide ratios from a SILAC quantitative proteomics experiment take on numbers in the theoretical, noninclusive range of zero to infinity represented by r_i where i is an enumeration of the ratios, taking values one to n . Without loss of generality, we will always consider the peptide ratios as being in the form of heavy over light. The number n denotes the total number of peptide ratios obtained in the experiment. After protein mapping, it is possible to assign each peptide ratio a protein identification number p_i between one and m where m is the total number of proteins identifiable by the peptides quantitated in the experiment. The number p_i is an enumeration of the identifiable proteins in some order such as by the protein's accession number. Together the values r_i and p_i allow us to identify each of the peptide ratios and to which protein they are matched.

Although the distribution of the peptide ratios $\{r_i\}_{i=1}^n$ is bounded to the left by zero, it has often been observed that the distribution of the logarithm of r_i : $l_i = \log(r_i)$, usually takes on a bell-shaped curve [24,27,28]. Under a 1:1 mix between heavy and light cells sampled from the same populations, we would expect the distribution to be centered at zero because the logarithm of one is zero and the abundances of proteins between the two populations should theoretically be the same [27]. Because of these properties, it is often more convenient and meaningful to work with the peptide log-ratios l_i instead of r_i .

2.1. Experimental Dataset

For the analysis of the distributions of peptide ratios, we utilise the data set from a published subcellular SILAC study comparing four fractions from prostate cancer PC-3 cells expressing the protein Polymerase I and Transcript Release Factor (PTRF) tagged with Green Fluorescent Protein (GFP), or GFP alone as control [3].

2.2. One sample t-tests

The one sample t-test for each protein j , is evaluated by first evaluating the mean over all peptide log-ratios matched to the protein numbered j : $L_j = \sum_{\{l | p_i = j\}} l_i / n_j$ where $n_j = \#\{l_i | p_i = j\}$ is the number of peptides matched to protein j . The protein log-ratios L_j are then divided by the standard error term $s_j / \sqrt{n_j}$ where $s_j = \sqrt{\sum_{\{l_i | p_i = j\}} (L_j - l_i)^2 / (n_j - 1)}$ is the standard deviation of the peptide log-ratio subsample $\{l_i | p_i = j\}$ which yields the test statistic

$$T_j = \frac{L_j}{s_j / \sqrt{n_j}}$$

Under the null hypothesis that there is no change in abundance of the protein between the two populations and under the assumption that the subsample $\{l_i | p_i = j\}$ is normal, the p-value of T_j can be found using the Student-t distribution [29]. It must also be noted that the t-test can only be applied to proteins that are quantitated by two or more peptides because of the necessity to estimate the standard deviation of the log-ratio subsamples.

2.3. Robust z-tests

To conduct a robust z-test, each of the proteins j are quantitated by taking the median over all peptide log-ratios matched to the protein numbered j : $M_j = \text{Median}(\{l_i | p_i = j\})$. An overall median of the protein log-ratios sample: $\{M_j\}_{j=1}^m : B_0$ alongside with the 15.87 and 84.13 percentiles: $P_{15.87}$ and $P_{84.13}$ can also be calculated. Assuming that the population of protein log-ratios has a Gaussian distribution, the differences $P_{50} - P_{15.87}$ and $P_{84.13} - P_{50}$ are both approximately equal to one standard deviation, and P_{50} is equal to the mean. Using these facts, the z-statistic

$$Z_j = \frac{(M_j - B_0)}{(P_{84.13} - P_{15.87}) / 2}$$

has a standard Gaussian distribution and can be used to evaluate the p-value of each of the proteins.

As an alternative, [25] specifies the test statistic

$$Z_j = \frac{M_j - P_{50}}{P_{84.13} - P_{50}}$$

for testing the null hypothesis of no difference in abundance versus the alternative hypothesis of greater abundance in the heavy population compared to the light population.

2.4. Wilcoxon signed-rank tests

Another robust hypothesis testing method is the Wilcoxon

signed-rank test [30]. For proteins, the Wilcoxon signed-rank test is used to test the hypothesis that the abundance is the same between the two cell populations against the hypothesis that the abundance being different. The p-value of the test is evaluated using the signed ranked peptide log-ratio subsample $\{l_i | p_i = j\}$ for each protein j through the Wilcoxon rank distribution or a normal approximation.

2.5. Permutation tests

We now present a permutation test procedure formulated in the tradition of [31] whereby we resample with replacement from the empirical data in order to evaluate the significance of a test statistic. In the context of bioinformatics, permutation testing has been applied to microarray gene expression studies [32,33], gene ontology and network analyses [34,35], and in proteomics specifically, permutation testing has been applied to false discovery rate control [36,37] and to biomarker discovery [38,39]. In the case of a SILAC protein expression experiment, we want to know the significance of the difference in the abundance of a protein between the heavy and light cell populations. The difference in abundance for a protein j can be measured as some function of the peptide log-ratio subsample $\{l_i | p_i = j\}$ such as the mean L_j or the median M_j . As an example, we will only consider L_j in following description of the method.

Under the assumption that the true population distribution of peptide log-ratios for the experiment resembles the sample distribution of the peptide log-ratios [40], we can calculate conservative estimates for the p-values of each protein j using the algorithm in (Figure 1) [41,42] for the null hypothesis that the protein is differently regulated between the two cell populations. Although there are no direct permutation of the protein labels p_i , the random sampling without replacement serves the practical purpose of permuting the class labels by obtaining n_j peptide log-ratios to label as protein j . The algorithm allows for working with data that may or may not need normalization such as 1:1 mixes and other mixing proportions respectively as well as allowing the replacement of the mean protein log-ratio L_j for any other test statistic of interest such as the median M_j .

It is important to note that, unlike the other hypothesis testing procedures, a p-value, obtained from (Figure 1) is an estimate of the true p-value for protein j with respect to the null hypothesis. In the case of a protein with n_j of one, the estimate p-value is the same each time the algorithm is run since the test computes the exact probability of obtaining the test statistic of L_j or more extreme under the null hypothesis in a similar fashion to Fisher's exact test [43]. In the cases where the proteins have subsample sizes n_j greater than one, we cannot calculate the exact p-value since it would require the exhaustive calculation of test statistics for all $n! / ((n - n_j)! n_j!)$ possible subsamples from $\{l_i\}_{i=1}^n$ of size n_j . This number grows rapidly and can become computationally infeasible for the sizes of most SILAC experiments. Instead of exhaustive computations over all possible subsamples, we

1. Set S as the sample of all peptide log-ratios $\{l_i\}_{i=1}^n$ from the experiment if the data does not require normalization.
 - a. If the data requires normalization, set S as the normalized sample $\{l_i - M\}_{i=1}^n$ or $\{l_i - \bar{l}\}_{i=1}^n$ where $M = \text{median}(\{l_i\}_{i=1}^n)$ and $\bar{l} = \sum_{i=1}^n \frac{l_i}{n}$ to normalize by the median or the mean respectively.
2. For protein j in the number of proteins one to m :
 - a. Calculate the number of peptides $n_j = \#\{l_i | p_i = j\}$ in the subsmaple of protein j .
 - b. Calculate the protein log-ratio $L_j = \sum_{\{l_i | p_i = j\}} \frac{l_i}{n_j}$ for protein j .
 - c. If $n_j = 1$:
 - i. Estimate the p-value of protein j by evaluating $p\text{-value}_j = \frac{\sum_{i=1}^n I\{|L_j| > |l_j|\} + 1}{n + 1}$ where $I\{|L_j| > |l_j|\} = 1$ if $|L_j| > |l_j|$ is true and $I\{|L_j| > |l_j|\} = 0$ if $|L_j| > |l_j|$ is false.
 - d. If $n_j > 1$:
 - i. For b from one to B :
 - A. Randomly sample without replacement n_j numbers between one and n to obtain the set $\{k_a\}_{a=1}^{n_j}$ where k_a is a number between one and n .
 - B. For the random peptide log-ratio subsample $S_b = \{k_a\}_{a=1}^{n_j}$ using the random sample $\{k_a\}_{a=1}^{n_j}$.
 - C. Compute the protein log-ratio of the random subsample by evaluating $L'_b = \sum_{a=1}^{n_j} \frac{k_a}{n_j}$.
 - ii. Estimate the p-value by evaluating $p\text{-value}_j = \frac{\sum_{b=1}^B I\{|L_j| > |L'_b|\} + 1}{B + 1}$ where $I\{|L_j| > |L'_b|\} = 1$ if $|L_j| > |L'_b|$ is true and $I\{|L_j| > |L'_b|\} = 0$ if $|L_j| > |L'_b|$ is false.

Figure 1. *Permutation testing p-value.* Algorithm for assessing the significance of protein log-ratios in SILAC expression experiments.

sample B of these subsamples with replacement and estimate the p-value based on these subsample test statistics instead. It is known that with increasing values of B the estimates $p\text{-value}_j$ approach their true values [41] and so B can be set based on obtaining a level of accuracy for $p\text{-value}_j$ or based on computational limitations.

2.6. Simulation setup

Assuming that the distribution of the peptide log-ratio sample $\{l_i\}_{i=1}^n$ is bell-shaped and has a mean of zero for a 1:1 mix, we can model the distribution of the peptide log-ratios of a quantitative proteomics experiment as a mixture of three bell-shaped distribution functions [44]. The three distributions must consist of a central, zero mean component representing the distribution of peptides which have no meaningful changes in abundance, a distribution with a negative mean which represents negatively changed peptides, and a distribution with positive mean which represents posi-

tively changed peptides. In the simplest case, a Gaussian distribution can be used to model each of the bell-shaped components of the log-ratio data [45]. A hypothetical protein log-ratio distribution under this model can be seen in (Figure 2).

The density of log-ratios composes a mixture of three Gaussian distribution functions can be expressed as

$$f(l) = \pi_1 f_G(l; \mu_1, \sigma_1^2) + \pi_2 f_G(l; \mu_2, \sigma_2^2) + \pi_3 f_G(l; \mu_3, \sigma_3^2)$$

whereby the parameters π_1 , π_2 and π_3 , each taking values between zero and one, denotes the mixing proportion of each of three components respectively, and $\pi_g f_G(l; \mu_g, \sigma_g^2)$ is the Gaussian density of the form

$$f_G(l; \mu_g, \sigma_g^2) = \frac{1}{\sqrt{2\pi\sigma_g^2}} e^{-\frac{(l-\mu_g)^2}{2\sigma_g^2}}$$

where μ_g and σ_g^2 are the mean and variance of component g respectively, and g is equal to one, two or three.

Of course the underlying distribution which models every experiment is different and will result in a different set of

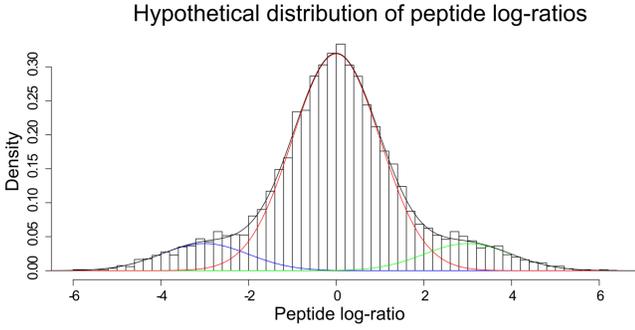


Figure 2. Hypothetical distribution of peptide log-ratios using a mixture of three Gaussian distributions. The black curve indicates the distribution of all peptide log-ratios in a protein expression experiment and the histogram is a sample of 10000 log-ratios from this distribution. The red curve indicates the distribution of all peptides that were unchanged in abundance, the blue curve indicates the distribution of negatively changed peptides, and the green curve indicates distribution of positively changed peptides. The black curve can be seen as the sum of the red, blue and green curves.

parameters for each of the mixture distribution components. We concentrate on a family of mixture distributions which we believe are representative of the data from many protein expression experiments. We fix the unchanged peptides distribution as a Gaussian distribution with a mean of zero and standard deviation of one. The positive and negative distributions are set to be symmetrical whereby the positive distribution has a mean of $\mu > 0$ and the negative distribution has a mean of $-\mu$. The variance and mixing proportions of the negative and positive distributions are equal and set to be σ^2 and π respectively. Therefore, we can write this family of peptide log-ratio distributions as

$$f(l) = (1 - 2\pi) f_G(l; 0, 1) + \pi f_G(l; -\mu, \sigma_2) + \pi f_G(l; \mu, \sigma_2)$$

whereby we can interpret 2π as the proportion of peptides that have changed in abundance in the experiment, $-\mu$ as the average negative change in the log-ratio and μ as the average positive change in the log-ratio for those peptides that have displayed a change in expression. Similar three component mixture models have been used in previous studies to model the unchanged, up-regulated and down-regulated SILAC log-ratios [44,46,47].

In order to cover as broad a range of distributions from this family as possible, we have chosen to simulate from the cases where μ is equal to 0.5, 1 or 2, σ^2 is equal to 0.25, 1 or 4, and 2π is equal to 0.1, 0.2 or 0.3. We believe these cases to be representative of the variety of relative variability and changes as well as the number of peptides that are differently expressed in an experiment. Simulating data from the experiment requires us to firstly determine the number of identified proteins m which we want to consider. For each of the m proteins j , we then randomly determine whether or not it is unchanged, negatively changed or positively changed with probabilities $1-2\pi$, π and π respectively. Once, we determine

behavior of the protein, we generate n_j peptide log-ratios for the protein from a Gaussian distribution with mean and variance corresponding to the behavior. The value of n_j for each protein comes from the geometric density function

$$f(n_j) = (1 - \phi)^{n_j - 1} \phi$$

where ϕ is a parameter which can take values between zero and one. We can justify the use of the geometric distribution because the minimum χ^2 -test for goodness-of-fit [48] indicates that the negative binomial distribution is appropriate for modeling the subsample sizes in the four data sets from [3] (with p-values of 1 in all cases) and the geometric distribution can be seen as a simple case of the negative binomial distribution. The full algorithm can be seen in (Figure 3). We choose to use the values $m = 500$ and $\phi = 0.05$ in all of our simulation scenarios since this results in an average total of 10000 peptide ratios which approximately corresponds to the values observed in the [3] data sets. Each set of parameters were simulated 10 times and some examples of simulated data can be seen in (Figure 4).

2.7. FDR, TPR and ACC

For each of the sets simulation parameters, we can evaluate the average estimated false discovery rate (FDR), true positive rate (TPR) and accuracy rate (ACC) [49] across the 10 repetitions. The FDR is a measure of the number of proteins incorrectly determined as changed, as a proportion of the number of proteins considered changed. The TPR is a measure of the number of changed proteins found as a proportion of the number of truly changed proteins, and the ACC is the total number of correctly classified proteins, both changed and unchanged as a proportion of the total number of proteins tested. All three values take on numbers from zero to one whereby a lower number is desirable for the FDR and a higher value indicates better performance for the TPR and ACC.

We can represent the true class labels of each of the m proteins in a simulated data set mathematically as $\{c_j\}_{j=1}^m$ where c_j is zero if the protein is unchanged and one if the protein is changed. Similarly, we can represent the result of the hypothesis tests as $\{t_j\}_{j=1}^m$ where t_j is zero if the p-value is greater or equal to 0.05 and one if the p-value is less than 0.05. Using this notation, the estimated FDR, TPR and ACC can be expressed as

$$\hat{FDR} = 1 - \frac{\sum_{\{A_{t_j=1}\}} I\{c_j = t_j\}}{\#\{j|t_j = 1\}}, \quad \hat{TPR} = \frac{\sum_{\{A_{c_j=1}\}} I\{c_j = t_j\}}{\#\{j|c_j = 1\}}, \quad \hat{ACC} = \frac{\sum_{j=1}^m I\{c_j = t_j\}}{m}$$

respectively where $I\{c_j = t_j\} = 1$ if $c_j = t_j$ is true and $I\{c_j = t_j\} = 0$ if $c_j \neq t_j$ is false. We can denote the average estimated FDR, TPR and ACC as \overline{FDR} , \overline{TPR} and \overline{ACC} respectively where they are each evaluated by averaging over the 10 repetitions of each set of simulations. These average measures allow us to comment on the performance of each of the hypothesis tests across the 27 different simulation scenarios when applied using the $\alpha = 0.05$ level of significance.

1. Set the proportion of changed proteins 2π .
2. Set the mean deviation in the expression of peptide log-ratios μ .
3. Set the variance of the changed log-ratio distributions σ^2 .
4. Set the parameter ϕ for generating the size of each protein subsample.
5. Set the number of proteins to simulate m .
6. For protein j in the set of proteins from one to m :
 - a. Randomly generate a protein subsample of size n_j from the geometric distribution $f(n_j) = (1-\phi)^{(n_j-1)} \phi$.
 - b. Randomly Draw R from a uniform distribution in the interval of zero and one.
 - c. If $0 \leq R \leq 1 - 2\pi$:
 - i. Randomly generate n_j peptide log-ratios from the density $f_G(l; 0, 1)$.
 - d. If $1 - 2\pi < R \leq 1 - \pi$:
 - i. Randomly generate n_j peptide log-ratios from the density $f_G(l; -\mu, \sigma^2)$.
 - e. If $1 - \pi < R \leq 1$:
 - i. Randomly generate n_j peptide log-ratios from the density $f_G(l; \mu, \sigma^2)$.

Figure 3. Peptide log-ratio simulation algorithm. Algorithm for simulating peptide log-ratio data based upon theoretical Gaussian mixture models.

2.8. AUC statistic

The area under the curve (AUC) measurement of the receiver operating characteristic (ROC) curve [50] was also calculated for each simulation scenario. The ROC curve plots the performance of a method in terms of the estimated false positive rate (FPR) and the estimated true positive rate achieved for each level of α used. The FPR can be interpreted as the number of truly unchanged proteins incorrectly classified as being changed as a proportion of the number of truly unchanged proteins. Like the other measurements, the FPR can take values between zero and one whereby lower values are desirable.

While the FPR and TPR are estimations of the probability of a type-I error and the power at every threshold level respectively, the AUC under the ROC curve is an average measure of performance of a method across all possible thresholds for a data set. Hence the AUC measurement has been widely adopted as a summary of performance in machine learning and proteomics applications [49,51–53].

We calculate the AUC for each method on each repetition of each set of simulation parameters. The average AUC measurement $\overline{\text{AUC}}$ for each testing method is calculated by averaging over the 10 repetitions of each simulation scenario. An example of ROC curves for an instance of the simulated data can be seen in (Figure 5). The average AUC, combined with the $\overline{\text{FDR}}$, $\overline{\text{TPR}}$ and $\overline{\text{ACC}}$ can give an indication of the relative performance of the permutation test against the competing testing methods.

3. Results

The normality assumption is necessary in guaranteeing accurate inferential results in many hypothesis tests. For this reason, the bell-shape of the distribution of the peptide log-ratios $\{l\}_{i=1}^n$ makes the assumption of a Gaussian distribution very tempting. It has been noted in the past that a Gaussian distribution is often not a good fit to actual peptide log-ratio data. We investigated this through observing histograms of the peptide-log ratios from our four subcellular fractions

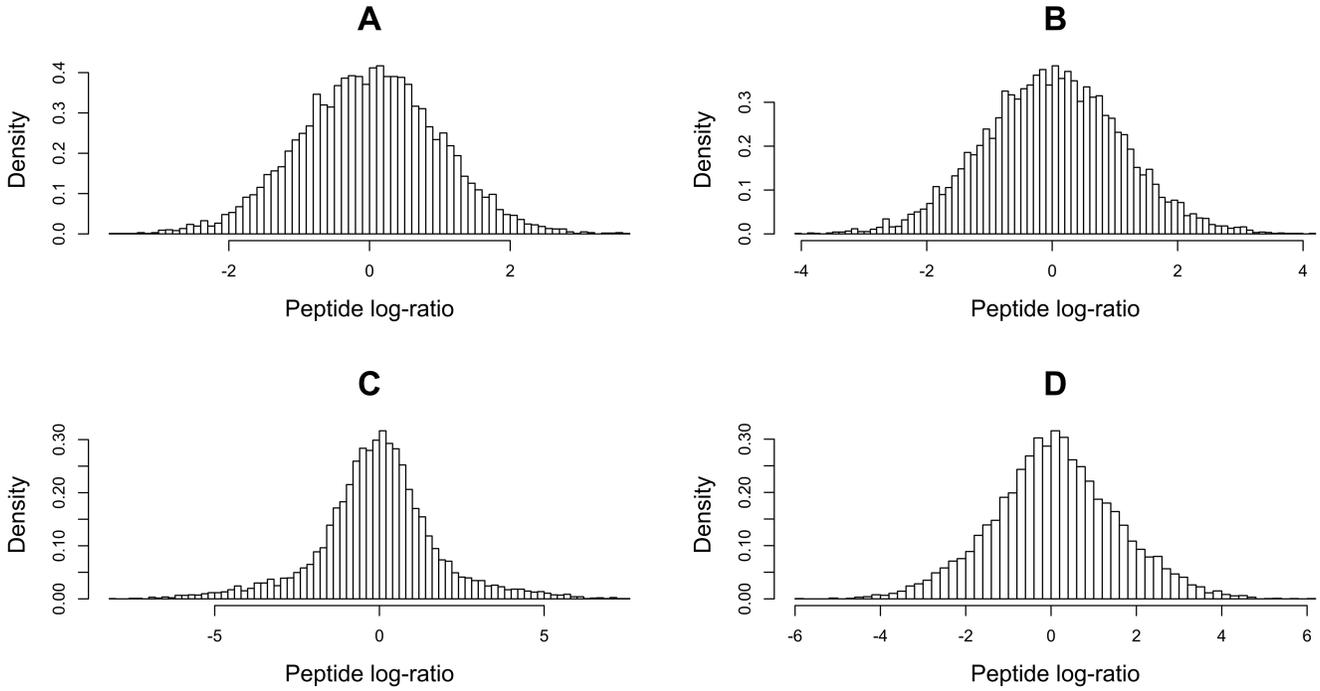


Figure 4. Histograms of simulated peptide log-ratios under different scenarios. Each of the histograms are an example of the simulated data of the peptide log-ratios from a protein expression experiment. The set of simulation parameters in each case are A: $2\pi = 0.1$, $\mu = 0.5$ and $\sigma^2 = 0.25$, B: $2\pi = 0.2$, $\mu = 0.1$ and $\sigma^2 = 1$, C: $2\pi = 0.3$, $\mu = 2$ and $\sigma^2 = 4$, and D: $2\pi = 0.3$, $\mu = 2$ and $\sigma^2 = 1$. We can see that changing the parameters can alter the peak, spread and shape of the tails of the distributions in each case.

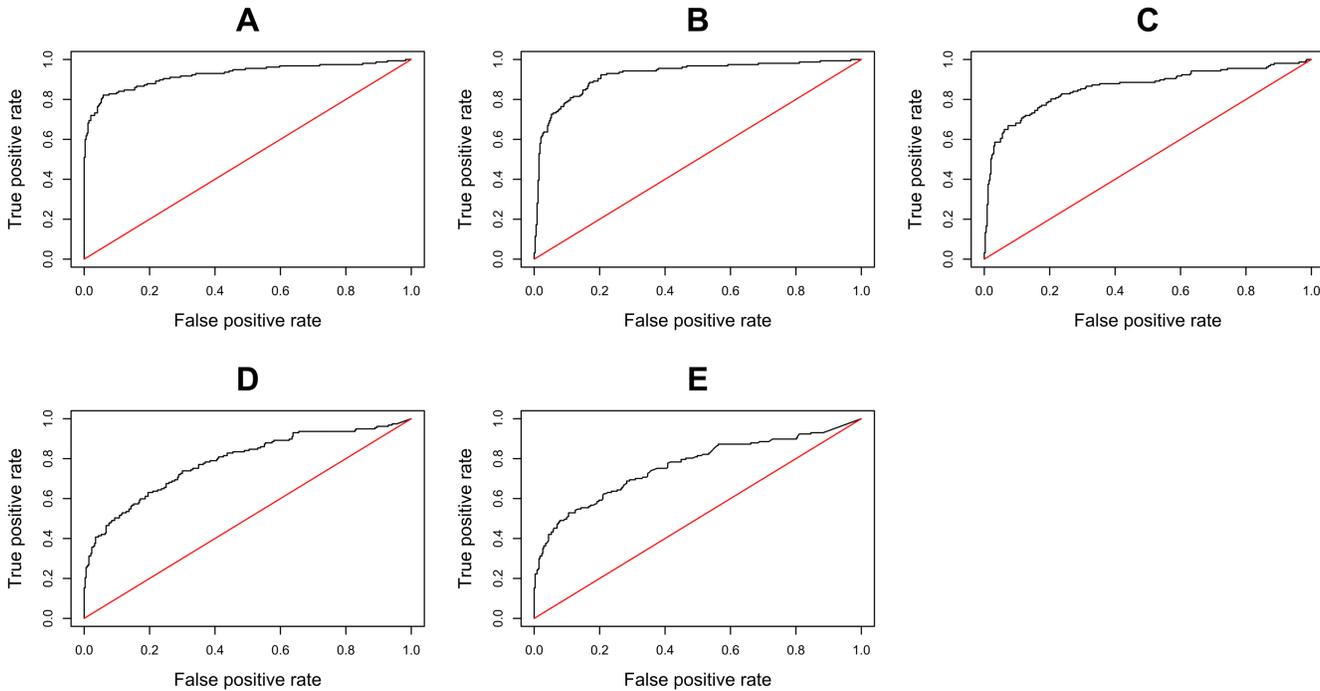


Figure 5. ROC curves for a simulation using the parameters $2\pi = 0.3$, $\mu = 2$ and $\sigma^2 = 4$. The ROC curve for each of the method is plotted as the black curve and the red line indicate the ROC curve for random guessing where TPR=FPR. The methods (with the AUC achieved over the data set) are A: permutation test (0.9252), B: z-test (0.9220), C: robust z-test (0.8600), D: t-test (0.7857) and E: Wilcoxon signed-rank test (0.7621).

data sets (Figure 6).

3.1. One-sample t-tests

Although the histograms are not Gaussian, traditional insight about the t-test often suggest that if the sample distribution does not deviate too greatly from the Gaussian distribution, a t-test may still be robust enough for use [56,57]. It has been shown that the t-test may still be applied when the data comes from a symmetric distribution which can often be the case when observing peptide log-ratio distributions over all proteins [58].

When we observe the largest peptide log-ratio subsamples from the four subcellular quantitations from [3] (Figure 7), we notice that not only is there strong deviations from symmetry in most cases but also multimodalism which severely violates the normality assumption. These results obtained from the largest subsamples from each of the four quantitations implies to us that it is difficult to assume normality for the peptide log-ratio populations for each of the quantitated proteins and thus invalidates the use of the one-sample t-test.

3.2. One-sample z-tests and robust z-tests

There is a similar problem with violations of the normality assumption in the use of the z-test whereby a Gaussian distribution is fitted to the protein log-ratio sample $\{L_j\}_{j=1}^m$ and the tails of the fitted distribution are used to calculate the p-value of each protein [60]. We can see the deviations from normality in the protein log-ratio histograms of the four subcellular quantitations from [3] (Figure 8). Because of the lack of fit of the Gaussian distribution to the protein log-ratio samples, the p-values calculated using fitted a curve will not accurately represent the true significance of each of the proteins.

There have been propositions for the use of more robust means of hypothesis testing. One such method is the robust z-test which uses the outlier-insensitive properties of the percentiles of a distribution rather than the mean and standard deviation which are known to be strongly influenced by extreme observations.

Under the assumption of Gaussian protein log-ratios, this z-statistic also has a standard Gaussian distribution and the p-value for each protein can be calculated by evaluating the tail probability $P(z > Z_j)$ where z is a standard Gaussian random variable. However, the method does not address the problem of deviations from normality in protein log-ratio distributions.

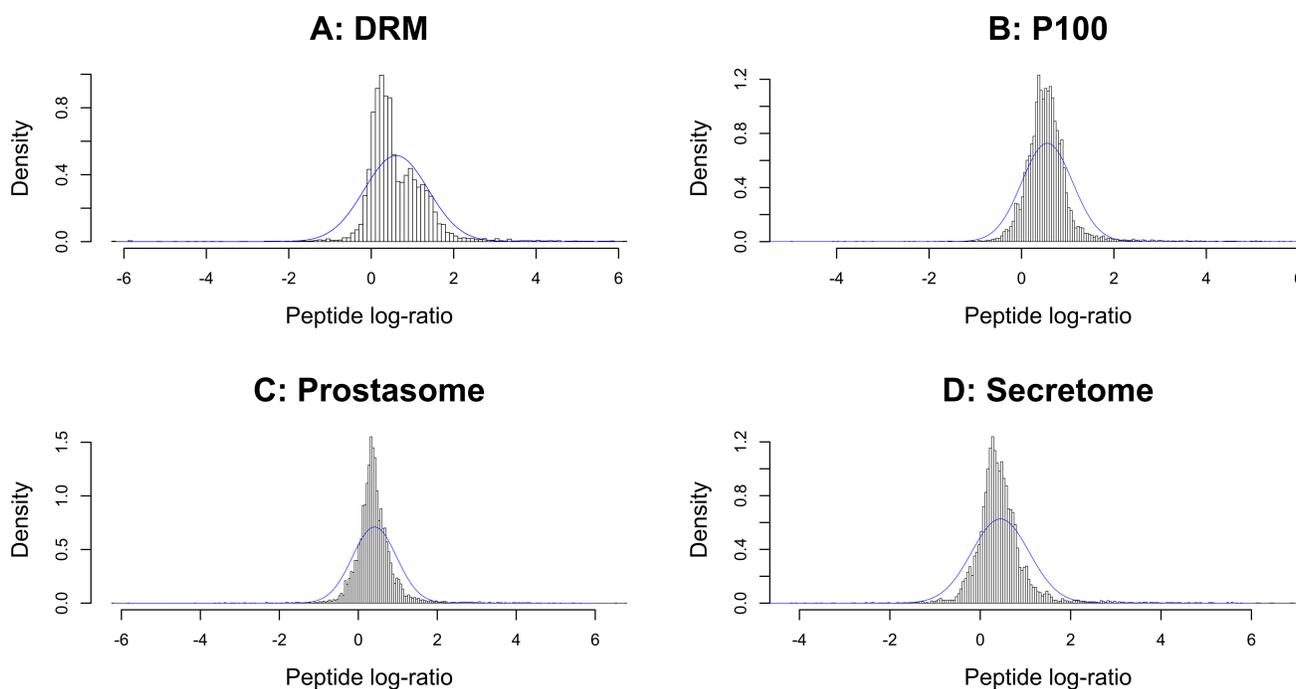


Figure 6. Histograms and fitted Gaussian distributions to peptide log-ratios of subcellular quantitations. The experimental data comes from the four subcellular proteome quantitations for the effect of Polymerase I and Transcript Release Factor (PTRF) expression in human prostate carcinoma PC-3 cells. The PTRF expressing cells were labeled with heavy amino acids and the cells without PTRF expression were labeled with light amino acids. The four subcellular fractionals (with number of peptides quantitated in brackets) were A: Detergent-Resistant Membrane (7338), B: Total membrane P100 fraction as described in [54] (10271), C: Prostasome (8682) and D: Secretome (9653). The Lilliefors test for normality p-values [55] for each sample was less than 2.2×10^{-16} indicating that the four sets of peptide log-ratios are severely non-Gaussian.

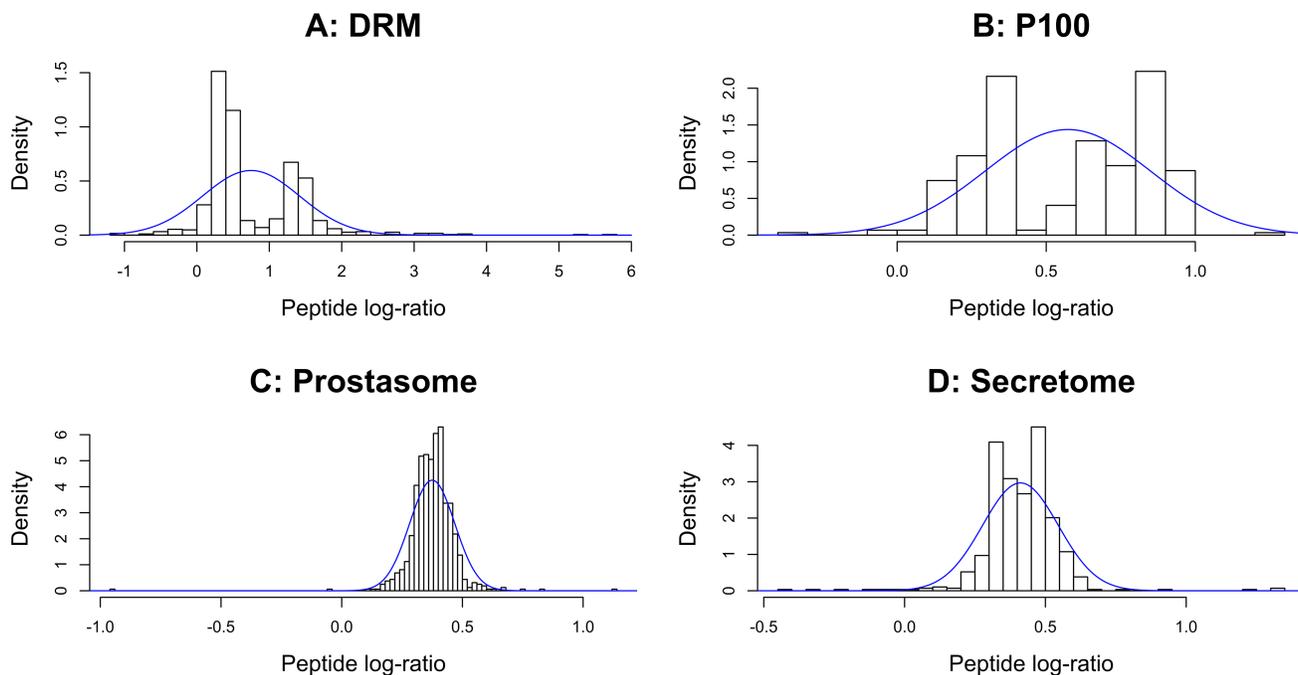


Figure 7. Histograms and fitted Gaussian distributions to peptide log-ratios of most common proteins. We plot the histograms of the peptide log-ratio subsample for the most commonly quantitated proteins of each of the quantitated fractions from [3]. The four most common proteins (with the number of peptides in the subsample in brackets) are A: Plectin-1 (929), B: Nucleophosmin (296), C: Thrombospondin-1 (802) and D: Thrombospondin-1 (577) coming from the DRM, P100, Prostatome and Secretome fractions respectively. The p-values for the D’Agostino test for skewness [59] for each subsample is A: less than 2.2×10^{-16} , B: 0.2656, C: less than 2.2×10^{-16} and D: 0.01957.

In order to assess the effects of violations of the normality assumption, we calculate the relative error between the true right-tail probability of a log-ratio of 3 in various Student-t distributions as compared to the right-tail probability calculated using the robust z-test as a ratio of the true probability. It is known that the Student-t distributions have heavier tails than the Gaussian distribution and asymptotically approaches the Gaussian distribution as the number of degrees of freedom approaches infinity [29]. We can see from (Figure 9) that the robust z-test would tend to underestimate the p-value of a protein log-ratio of 3 when the distribution of protein log-ratios has heavier than Gaussian tails. Because of the discrepancies in the p-values, the robust z-test would tend to have a higher proportion of false positives when the protein log-ratios come from distribution with heavier than Gaussian tails.

3.3. Wilcoxon signed-rank tests

The Wilcoxon test is a nonparametric hypothesis testing procedure that does not depend on the distribution of the underlying population. The procedure is therefore robust to deviations from normality and can therefore be implemented to sample data regardless of the underlying population distribution. The cost to this robustness is that the test lacks the power to detect large deviations from the null hypothesis in small samples.

To illustrate the lack of power of the test, suppose that we

only observe positive peptide log-ratios in the subsample $\{l_i | p_i = j\}$ for a protein j . The p-value of the Wilcoxon signed rank test at various number of peptides $\#\{l_i | p_i = j\}$ can be seen in (Figure 10) showing that a protein needs to be quantitated by at least six peptides in order to be deemed significantly different within the two cell populations at the five percent level. Proteins which are quantitated only by positive peptide log-ratios are a best case scenario and thus, for proteins that are quantitated by both positive and negative peptide log-ratios, the number of peptides needed for a significant result would be greater than six. The lack of power makes it impossible to find significant differences for rare proteins between the two cell populations and therefore diminishes the usefulness of the procedure.

3.4. Permutation tests

The permutation test proposed here is both distribution free and more powerful than the Wilcoxon test in small samples. The lack of distributional assumptions allows the permutation tests to be applicable to SILAC log-ratio data which as we have shown to not meet the required assumptions for conventional hypothesis tests. Additionally, the permutation test appears to be better suited than the other procedures, especially for assessing the significance of proteins that are quantitated by small numbers of peptides.

We now report on the simulation study conducted to assess the performance of the permutation test against the t-

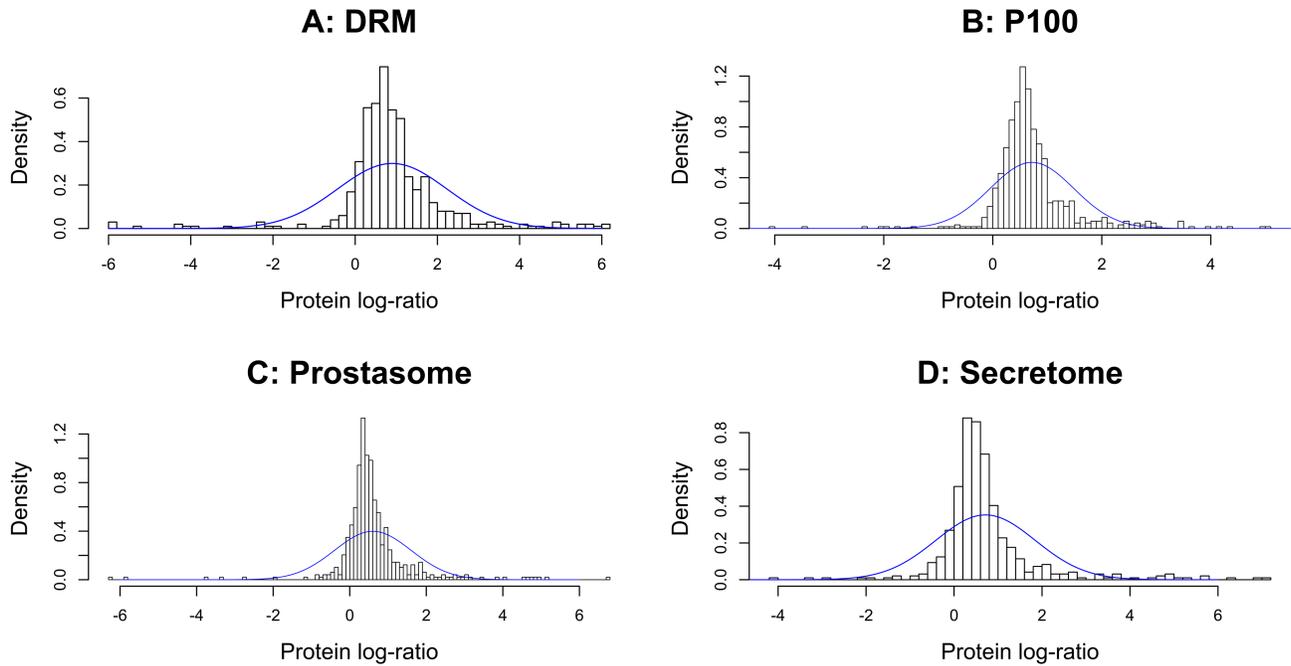


Figure 8. Histograms and fitted Gaussian distributions to protein log-ratios of subcellular quantitations. The four subcellular fractions (with number of protein quantitated in brackets) were A: Detergent-Resistant Membrane (504), B: P100 (691), C: Protasome (489) and D: Secretome (484). The Lilliefors test for normality p-values for each sample was less than 2.2×10^{-16} shows that the protein log-ratio samples are significantly non-Gaussian in distribution.

test, z-test, robust z-test and Wilcoxon signed-rank test across the 27 simulation scenarios. The performance for each method in each scenario is measured by the FDR, TPR, ACC and AUC averaged over the 10 repetitions.

3.5. Evaluating competing methodologies

We firstly assess the performance of each of the methods as applied with the significance level of $\alpha = 0.05$. Since each of the hypothesis testing methods results in a p-value for each protein, we can use the rule that if the p-value is less than 0.05, we consider that the protein is differently abundant between our two samples and if the p-value is greater or equal to 0.05 then we consider that the protein is unchanged. For the t-test, a p-value of 1 is given to any protein that it cannot quantitate since that protein would generally undergo no further assessment and therefore is vacuously deemed unchanged. The number of permutations B is set to 10000 which have been shown to result in sufficiently accurate p-values when tested against a Gaussian null distribution. Results of the FDR, TPR, ACC and AUC across the 27 different simulation scenarios can be found respectively in (Table S1), (Table S2), (Table S3) and (Table S4) in Supplementary Section 1.

To facilitate comparison, we also ranked the five methods for each of the set of parameters, where one represents the method that performed the worst in a scenario and five represents the best method. Ties are each given fractional ranks equal to the average of the tied ranks. The ranked results for

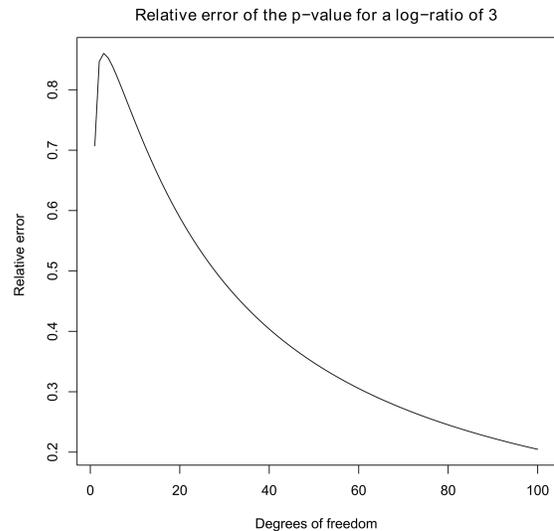


Figure 9. Relative errors between the true p-value and robust z-test p-value for a log-ratio of 3. The relative error at is calculated as $(P_t - P_z)/P_t$ where $P_t = P(t < (3 - P_{50}) / (P_{84.13} - P_{50}))$, $P_z = P(z < (3 - P_{50}) / (P_{84.13} - P_{50}))$, t is a Student-t random variable with degree of freedom ν , z is a standard Gaussian random variable and the percentiles come from a Student-t distribution with degrees of freedom ν . The relative error for degrees of freedom of one to 100 are graphed showing that the true p-value is always greater than those calculated using the robust z-test. The relative error is decreasing with increasing degrees of freedom because the shapes of the Student-t distribution approaches normality as the degrees of freedom approaches infinity.

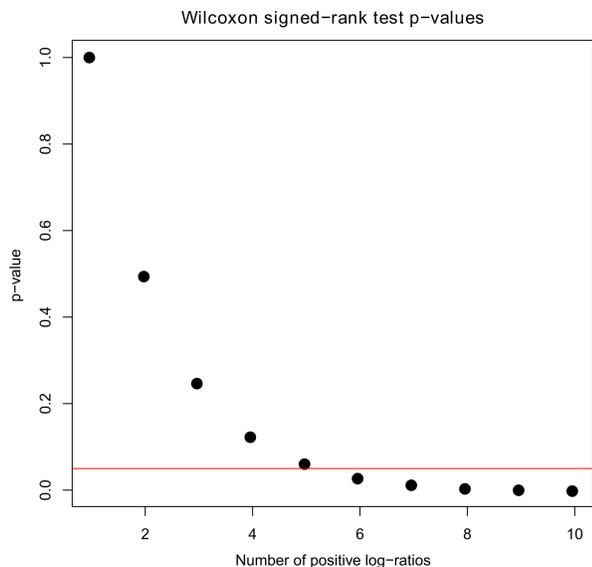


Figure 10. *p*-values of a Wilcoxon signed-rank test for various numbers of positive sample observations. The graph shows the *p*-values of a Wilcoxon signed-rank test for various sized samples of positive values. The red line indicates the five percent level of significance and indicates that we need at least six observations in order to obtain a *p*-value less than the significance level. The graph also shows that the *p*-value is decreasing with respects to larger numbers of observations, indicating that the more positive peptide log-ratios observed, the greater the probability that there is a difference between the protein abundances of the heavy and light cell populations.

the \overline{FDR} , \overline{TPR} and \overline{ACC} and \overline{AUC} across the 27 different simulation scenarios can be found respectively in (Table S5), (Table S6), (Table S7) and (Table S8) in Supplementary Section 1.

Using the rankings, we can calculate the total of the ranks over each of the measurements. These totals give us an indication of the relative performance for each of the testing procedures over all simulation scenarios. The total ranks for each of the six testing methods over the four measurements can be found in (Table 1).

From the total ranks results we can see that the permutation test performed the best in terms of \overline{FDR} and \overline{ACC} , second best in terms of \overline{AUC} , and third best in terms of \overline{TPR} . These results are subjected to variability due to the random-

ized process of the simulation study. It is therefore also interesting to observe the statistical significance of each of the difference between the ranks of the permutation test and those of the competing methods. The paired-sign test [61] is applied to the rank data for each of the four measurements in order to perform this assessment. The *p*-values for each of the tests can be found in (Table 2).

The results of the hypothesis tests suggest that at the 0.05 level, there is no significant difference between the ranks of the permutation test and the z-test over both the \overline{TPR} and \overline{AUC} . Therefore, there is not enough evidence to suggest that the z-test outperforms the permutation test in those two measurements.

4. Discussion

The tabulated results show that the permutation test can be seen as being the best performing method or tied for best performing method over the average \overline{FDR} , \overline{ACC} and \overline{AUC} measurements. This indicates that over a variety of possible simulation scenarios, the permutation test can be seen to either equal or outperform the z-test, robust z-test, t-test and Wilcoxon signed-rank test. The performance over the \overline{FDR} and \overline{ACC} indicates that the permutation test makes the fewest false identification of changed proteins as well as the fewest mislabeling of proteins across the different simulation parameter sets when applying a $\alpha = 0.05$ threshold. The result from the \overline{AUC} shows that the permutation test is also the best test to use with any arbitrary threshold as well, indicating that it is not only optimal at the 0.05 level.

The permutation test is only significantly bettered by the robust z-test in terms of ranks of \overline{TPR} . Observing (Table 2), we note that although the permutation test is often beaten by the robust z-test, the absolute value of the \overline{TPR} in many scenarios are very small. We also note that the robust z-test performs the worst in terms of \overline{FDR} and thus its strong performance can be attributed to the method's inaccuracy.

We also note that the simulation setup used is favorable to the tests which rely on the assumption of normality. This is because the Gaussian mixture models have tails of similar shapes to single Gaussian distributions which favors the z-test and the robust z-test. The sampling of the peptide log-ratios for each protein subsample from a single Gaussian distribution strongly favors the t-test since the assumption of

Table 1. The total ranks for each of the hypothesis testing procedures over the \overline{FDR} , \overline{TPR} , \overline{ACC} and \overline{AUC} measurements.

Measurement — Test	Permutation test	z-test	Robust z-test	t-test	Wilcoxon signed-rank test
\overline{FDR}	124	65	35	107	74
\overline{TPR}	75	90	134	35	71
\overline{ACC}	121	67	47	84.5	85.5
\overline{AUC}	108	115	77	67	38

Table 2. Paired-sign p-values for the differences in ranks over the \bar{FDR} , \bar{TPR} , \bar{ACC} and \bar{AUC} measurements.

Measurement — Test	z-test	Robust z-test	t-test	Wilcoxon signed-rank test
\bar{FDR}	4.172×10^{-7}	4.172×10^{-7}	5.925×10^{-3}	4.923×10^{-5}
\bar{TPR}	0.1221	1.49×10^{-8}	1.514×10^{-3}	0.2478
\bar{ACC}	5.648×10^{-6}	4.172×10^{-3}	5.925×10^{-3}	1.514×10^{-3}
\bar{AUC}	1	0.01916	4.923×10^{-5}	4.923×10^{-5}

normality within the subsamples are held. In order to test the performance of the permutation test under deviations from normality, we simulated two additional scenarios where we applied the Laplace distribution [62] and the Student-t distribution with $\nu = 3$ degrees of freedom. The results of these simulations can be found in Supplementary Section 2.

Analyses of these additional simulations show that the permutation test has equivalent performance in both the Gaussian and leptokurtic simulation scenarios. Our choices of possible distributions for modeling the changed peptides distributions are by no means exhaustive. Other examples of possible deviations from normality include modeling the log-ratios with Cauchy distribution mixtures [63], modeling the tails with generalized Pareto distributions [44] or using asymmetric up-regulated and down-regulated distributions.

4.1. False discovery rate mitigation

In the simulation scenarios, it is possible to calculate the estimated FDR in each case due to the knowledge of the class labels. However, in an experimental situation, where the true nature of each protein is not known, it is desirable to control the FDR at an acceptable level such that the FDR does not inflate the multiple testing error rate of any downstream analysis. Because the permutation test outputs a p-value, it is easy to perform FDR control either with techniques such as the Benjamini-Hochberg procedure [64] or with empirical Bayes methods [65,66].

In any case, it must be acknowledged that with stricter control of the FDR, there is a trade-off of a reduction in the TPR. Therefore, the cost of controlling the number of false discoveries made is that less of the truly changed proteins will be declared as having a significant change in abundance.

4.2. Other methodologies

Apart from the z-test, Robust z-test, t-test and Wilcoxon signed-rank test, there are many alternative methods available for quantitative proteomics experiments. Such methods include spectral counting methods and multiple experimental replicates methods such as those found in [67] and [22]. These methods, unlike the methods for ratios and log-ratios rely on greater volumes of data from multiple repeated

experiments or richer experimental data outputs. Such comprehensive data is unavailable in many experimental situations, so we do not discuss these methods further.

Other methods which only rely on single ratio data sets such as likelihood based approaches [44,68]. These methods were not used for comparison because they are often application specific and require statistical expertise which is not available to many researchers. For this reason, we only chose to discuss simple hypothesis testing methods such as permutation testing on log-ratios since this is interpretable and implementable by most researchers using data that is already available to them.

4.3. Broader applications

Although the discussion has focused on permutation testing of SILAC ratio data, as successfully implemented in [3], the permutation test can also be applied to data resulting from other relative quantitation methods such as isobaric tags for relative and absolute quantitation (iTRAQ) [69], isotope-code affinity tags (ICAT) [70], label-free quantitation [67] and tandem mass tags (TMT) [71]. Outside of quantitative proteomics, it is also possible to apply the permutation testing method to fold-change data from microarray gene expression experiments [33].

5. Concluding Remarks

The permutation test which we introduced is distribution free, applicable across a range of relative quantitation methods, and was found to be superior to the other tests across the spectrum of significance level through analyzing the AUC statistics. At the usual 0.05 level it was also found the permutation test generally performed better than the competing methods over the ACC, FDR and TPR criteria.

The advantage of the permutation testing procedure over more technical methods is that it is a simple hypothesis test that outputs a p-value which can be interpreted by scientists. This allows it to be controlled for FDR and inferred in the same way as the p-values of other simple tests.

The permutation method is implementable in any programming environment and an R implementation of the procedure is available upon request. We are currently assembling a server-based web interface to for the test with ease of

deployment in mind. This web interface will provide additional data visualization and summarization facilities along with the computation of the permutation p-values.

6. Supplementary material

Supplementary Material.pdf contains Supplementary Section 1 and Supplementary Section 2.

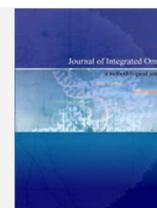
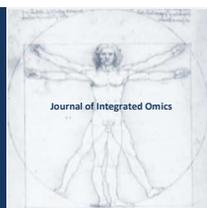
Acknowledgements

This work was supported by a University of Queensland summer research scholarship to HN, and The University of Queensland Diamantina Institute Fellowship to MH. MH is supported by an NHMRC Career Development Award No. 569512). We gratefully thank Kim-Anh Le Cao for her critique of the manuscript. We would also like to thank the anonymous reviewers for their helpful comments which lead to improvements in the paper.

References

- [1] S.-E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol Cell Proteomics* 1 (2002) 376–386.
- [2] J.S. Andersen, Y.W. Lam, A.K.L. Leung, S.-E. Ong, C.E. Lyon, A.I.L. & M. Mann, *Nature* 433 (2005) 77–83.
- [3] K.L. Inder, Y.Z. Zheng, M.J. Davis, H. Moon, D. Loo, H. Nguyen, J.A. Clements, R.G. Parton, L.J. Foster, M.M. Hill, *Mol Cell Proteomics* 11 (2012) M111.012245.
- [4] S.E. Ong, L.J. Foster, M. Mann, *Method Mol Cell Biol* 29 (2003) 124–130.
- [5] E. Milner, E. Barnea, I. Beer, and A. Admon, *Mol Cell Proteomics* 5 (2006) 357–365.
- [6] B. Soufi, C. Kumar, F. Gnad, M. Mann, I. Mijakovic, and B. Macek, *J Proteome Res* 9 (2010) 3638–3646.
- [7] A. Cuomo, T. Bonaldi, *Method Mol Cell Biol* 662 (2010) 59–78.
- [8] M.D. Sury, J.X. Chen, M. Selbach, *Mol Cell Proteomics* 9 (2010) 2173–2183.
- [9] W. Schütz, N. Hausmann, K. Krug, R. Hampp, B. Macek, *Plant Cell* 23 (2011) 1701–1705.
- [10] A. Gruhler, J.V. Olsen, S. Mohammed, P. Mortensen, N.J. Faergeman, M. Mann, O.N. Jensen, *Mol Cell Proteomics* 4 (2005) 310–327.
- [11] T. Geiger, J. Cox, P. Ostasiewicz, J.R. Wisniewski, M. Mann, *Nat Methods* 7 (2010) 383–385.
- [12] M. Mann, *Nat Rev Mol Cell Bio* 7 (2006) 952–958.
- [13] E.A. Kapp, F. Schütz, L.M. Connolly, J.A. Chake, J.E. Meza, C.A. Miller, D. Fenyo, J.K. Eng, J.N. Adkins, G.S. Omenn, R.J. Simpson, *Proteomics* 5 (2005) 3475–3490.
- [14] K.W. Lau, A.R. Jones, N. Swainston, J.A. Siepen, S.J. Hubbard, *Proteomics* 7 (2007) 2787–2799.
- [15] K.T. Rigbolt, B. Blagoev, *Method Mol Cell Biol* 658 (2010) 187–204.
- [16] H. Choi, A.I. Nesvizhskii, *J Proteome Res* 7 (2008) 47–50.
- [17] J. Graumann, N.C. Hubner, J.B. Kim, K. Ko, M. Moser, C. Kumar, J. Cox, H. Schöler, M. Mann, *Mol Cell Proteomics* 7 (2008) 672–683.
- [18] M. Ackermann, K. Strimmer, *BMC Bioinformatics* 10 (2009) 47.
- [19] D.W. Huang, B.T. Sherman, R.A. Lempicki, *Nucleic Acids Res* 37 (2009) 1–13.
- [20] S.Y. Rhee, V. Wood, K. Dolinski, S. Draghici, *Nat Rev Genet* 9 (2008) 509–515.
- [21] B. van Breukelen, H.W.P. van den Toorn, M.M. Drugan, and A.J.R. Heck, *Bioinformatics* 25 (2009) 1472–1473.
- [22] J.W.H. Wong, M.J. Sullivan, G. Cagne, *Brief Bioinform* 9 (2007) 156–165.
- [23] N.T. Seyfried, Y.M. Gozal, E.B. Dammer, Q. Xia, D.M. Duong, D. Cheng, J.J. Lah, A.I. Levey, and J. Peng, *Mol Cell Proteomics* 9 (2010) 705–718.
- [24] W.-T. Lin, W.-N. Hung, Y.-H. Yian, K.-P. Wu, C.-L. Han, Y.-R. Chen, Y.-J. Chen, T.-Y. Sung, and W.-L. Hsu, *J Proteome Res* 5 (2006) 2328–2338.
- [25] J. Cox, M. Mann, *Nat Biotechnol* 26 (2008) 1367–1372.
- [26] J. Schimmel, K.M. Larsen, I. Matic, M. van Hagen, J. Cox, M. Mann, J.S. Andersen, A.C.O. Vertegaal, *Mol Cell Proteomics* 7 (2008) 2107–2122.
- [27] S. ten Have, S. Boulon, Y. Ahmad, A.I. Lamond, *Proteomics* 11 (2011) 1153–1159.
- [28] P. Mertins, H.C. Eberl, J. Renkawitz, J.V. Olsen, M.L. Tremblay, M. Mann, A. Ullrich, H. Daub, *Mol Cell Proteomics* 7 (2008) 1763–1777.
- [29] Student, *Biometrika* 6 (1908) 1–25.
- [30] F. Wilcoxon, *Biometrics Bull* 1 (1945) 80–83.
- [31] M. Dwass, *Ann Math Stat* 28 (1957) 181–187.
- [32] S. Dudoit, Y.H. Yang, M.J. Callow, T.P. Speed, *Stat Sinica* 12 (2002) 111–139.
- [33] V.G. Tusher, R. Tibshirani, G. Chu, *P Natl Acad Sci USA* 98 (2001) 5116–5121.
- [34] F. Al-Shahrour, R. Díaz-Uriarte, J. Dopazo, *Bioinformatics* 20 (2004) 578–580.
- [35] J.J. Goeman, P. Bühlmann, *Bioinformatics* 23 (2007) 980–987.
- [36] C. Schaab, *Method Mol Cell Biol* 696 (2011) 41–57.
- [37] D.M. Walther, M. Mann, *Mol Cell Proteomics* 10 (2011) M110.004523.
- [38] S. Smit, M.J. van Breemen, H.C.J. Hoefsloot, A.K. Smilde, J.M.F.G. Aerts, C.G. de Koster, *Anal Chim Acta* 592 (2007) 210–217.
- [39] S. Smit, H.C.J. Hoefsloot, A.K. Smilde, *J Chromatogr B* 866 (2008) 77–88.
- [40] A.W. van der Vaart, *Asymtotic Statistics*, Cambridge University Press, 1998.
- [41] M.D. Ernst, *Stat Sci* 19 (2004) 676–685.
- [42] B. Phipson, G.K. Smyth, *Stat Appl Genet Mo B* 9 (2010) 39.
- [43] R.A. Fisher, *Statistical Methods for Research Workers*, Oliver and Boyd, 1958.
- [44] A.A. Margolin, S.-E. Ong, M. Schenone, R. Gould, S.L. Schreiber, S.A. Carr, T.R. Golub, *PLoS One* 4 (2009) e7454.

- [45] G.J. McLachlan, K.E. Basford, *Mixture Models: Inference and Applications to Clustering*, Marcel Dekker, 1988.
- [46] A. von Kriegsheim, D. Baiocchi, M. Birtwistle, D. Sumpton, W. Bienvenu, N. Morrice, K. Yamada, A. Lamond, G. Kalna, R. Orton, D. Gilbert, W. Kolch, *Nat Cell Biol* 11 (2009) 1458–1464.
- [47] S.-E. Ong, M. Schenone, A.A. Margolin, X. Li, K. Do, M.K. Doud, D.R. Mani, L. Kuai, X. Wang, J.L. Wood, N.J. Tolliday, A.N. Koehler, L.A. Marcaurrelle, T.R. Golub, R.J. Gould, S.L. Schreiber, S.A. Carr, *P Natl Acad Sci USA* 106 (2009) 4617–4622.
- [48] W.G. Cochran, *Ann Math Stat* 23 (1952) 315–345.
- [49] T. Fawcett, *Pattern Recogn Lett* 27 (2006) 861–874.
- [50] A.P. Bradley, *Pattern Recogn* 30 (1997) 1145–1159.
- [51] N. Chen, W. Sun, X. Deng, Y. Hao, X. Chen, B. Xing, W. Jia, J. Ma, H. Wei, Y. Zhu, X. Qian, Y. Jiang, F. He, *Proteomics* 8 (2008) 5108–5118.
- [52] J. Cui, X. Ma, L. Chen, J. Zhang, *BMC Bioinformatics* 12 (2011) 439.
- [53] S.R. Ramakrishnan, C. Vogel, T. Kwon, L.O. Penalva, E.M. Marcotte, D.P. Miranker, *Bioinformatics* 25 (2009) 2955–2961.
- [54] K.L. Inder, C. Lau, D. Loo, N. Chaudhary, A. Goodall, S. Martin, A. Jones, D. van der Hoeven, R.G. Parton, M.M. Hill, J.F. Hancock, *J Biol Chem* 284 (2009) 28410–28419.
- [55] H.W. Lilliefors, *J Am Stat Assoc* 62 (1967) 399–402.
- [56] J.V. Bradley, *Brit J Math Stat Psy* 31 (1978) 144–152.
- [57] H. Hotelling, in: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, 1961, .
- [58] B. Efron, *J Am Stat Assoc* 64 (1969) 1278–1302.
- [59] R.B. D’Agostino, *Biometrika* 57 (1970) 679–681.
- [60] J.-Y. Zhou, J. Hanfelt, J. Peng, *Proteom Clin Appl* 1 (2007) 1342–1350.
- [61] W.J. Dixon, A.M. Mood, *J Am Stat Assoc* 41 (1946) 557–566.
- [62] C. Forbes, M. Evans, N. Hastings, B. Peacock, *Statistical Distributions*, Wiley, 2011.
- [63] F.P. Breitwieser, A. Müller, L. Dayon, T. Köcher, A. Hainard, P. Pichler, U. Schmidt-Erfurth, G. Superti-Furga, J.-C. Sanchez, K. Mechtler, K.L. Bennett, and J. Colinge, *J Proteome Res* 10 (2011) 2758–2766.
- [64] Y. Benjamini, Y. Hochberg, *J Roy Stat Soc B* 57 (1995) 289–300.
- [65] B. Efron, *J Am Stat Assoc* 99 (2004) 96–104.
- [66] G.J. McLachlan, R.W. Bean, L.B.-T. Jones, *Bioinformatics* 22 (2006) 1608–1615.
- [67] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, *Anal Bioanal Chem* 389 (2007) 1017–1031.
- [68] C. Pan, G. Kora, W.H. McDonald, D.L. Tabb, N.C. VerBerkmoes, G.B. Hurst, D.A. Pelletier, N.F. Samatova, R.L. Hettich, *Anal Chem* 78 (2006) 7121–7131.
- [69] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlett-Jones, F. He, A. Jacobson, D.J. Pappin, *Mol Cell Proteomics* 3 (2004) 1154–1169.
- [70] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, *Nat Biotechnol* 17 (1999) 994–999.
- [71] A. Thompson, J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, and C. Hamon, *Anal Chem* 75 (2003) 1895–1904.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.102

Analysis of the rat primary hepatocyte nuclear proteome through sub-cellular fractionation.

Cliff Rowe, Roz E. Jenkins, Neil R. Kitteringham, B. Kevin Park, Christopher E.P. Goldring*

MRC Centre for Drug Safety Science, Department of Pharmacology & Therapeutics, Sherrington Building, Ashton Street, The University of Liverpool, L69 3GE.

Received: 09 August 2012 Accepted: 17 October 2012 Available Online: 02 November 2012

ABSTRACT

Characterising primary hepatocytes and their de-differentiation in culture is vital for the refinement of current culture techniques and for the development of new and improved in vitro hepatocyte models. We have performed multiplexed iTRAQ proteomics on whole cell preparations and further employed nuclear fractionation to expand the coverage of this important organelle. We identify many proteins that change in abundance during culture of rat hepatocytes for 48h and map their molecular functions. 431 proteins were identified and quantified in whole cell homogenates, mapping to 69 molecular functions using the PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system. In whole cell homogenates liver-associated functions, such as oxidoreductase activity, were enriched compared with the reference rat proteome dataset but some functions, such as transcriptional activity, were under-represented. Nuclear fractionation resulted in the identification of an additional 156 proteins which mapped to 31 molecular functions. These proteins included some associated with hepatic differentiation, such as HNF4alpha and CCAAT/enhancer-binding protein beta and others with less well-defined roles.

Hierarchical clustering of samples within each experiment showed segregation of fresh and cultured sample types and stringent statistical analysis demonstrated significant changes in 36% of proteins from the whole cell homogenates and 21% of proteins from the nuclear dataset (adjusted $p < 0.05$). The molecular functions of the changed proteins in each dataset are mapped. These datasets broaden our understanding of hepatocyte de-differentiation and will aid the identification of target pathways to attenuate de-differentiation in culture and maintain hepatocytes with a more relevant physiological phenotype.

Keywords: Differentiation; Hepatocyte; Nucleus; Proteomics.

Abbreviations:

ACN, acetonitrile; HBSS, Hanks balanced salt solution; iTRAQ, isobaric tags for relative and absolute quantification; LC-MS/MS, Liquid chromatography mass spectrometry/mass spectrometry; MMTS, Methyl methanethiosulfonate; PBS, Phosphate buffered saline; SDS, Sodium Dodecyl sulphate; TFA, trifluoroacetic acid; TEAB, triethylammonium bicarbonate; WEM, Williams' E Medium.

1. Introduction

Mammalian hepatocyte cultures represent an established and essential tool in drug development [1-5]. However, the relatively poor availability of human cells and their limited lifespan in culture means that animal cells, particularly rat

hepatocytes, are often employed. Although these cells have utility, there is an urgent need to develop new, robust, metabolically-competent human hepatocyte models with predictable availability for routine hepatotoxicity and

*Corresponding author: Christopher EP Goldring, E-mail Address: C.E.P.Goldring@liverpool.ac.uk; Postal Address: MRC Centre for Drug Safety Science, Department of Pharmacology & Therapeutics, Sherrington Building, Ashton Street, The University of Liverpool, United Kingdom, L69 3GE.

metabolism studies. Despite the continuous refinement of hepatocyte culture techniques, these cells undergo profound changes during culture with loss of liver function [6]. The underlying mechanisms are partially understood [7] but our knowledge concerning which proteins, and cellular functions change during culture remains incomplete. In a recent study we profiled the proteome of freshly-isolated hepatocytes and mapped temporal changes in primary culture that were concomitant with large declines in the abundance of liver specific mRNAs [6]. We postulated that significant changes in the nuclear proteome may be driving de-differentiation of hepatocytes in culture although there was limited coverage of this organelle in our dataset.

The nucleus of a cell orchestrates its phenotype through transcriptional control of specific gene programs. 'Master' transcription factors that determine cell specificity during embryonic development have been identified for many cell types [8] including hepatocytes [9], but alone these represent an incomplete picture of the complex networks of transcription factors, co-factors, chromatin remodelling proteins, RNA processing factors, and myriad accessory proteins that combine to direct cellular phenotype. Studies have been conducted previously to identify the proteins found in the nuclei of various cell types [10-14]. In the context of hepatocytes, a wider knowledge of the nucleus, and how its proteome changes in culture, should help develop better models of primary culture and assist in defining pathways amenable to manipulation. In this study, therefore, our aim was to further our understanding of hepatocyte de-differentiation through analysis of freshly-isolated and cultured hepatocyte whole cells and purified nuclei using iTRAQ proteomics and analysis of the molecular functions associated with the proteins that significantly change in abundance.

2. Material and Methods

2.1 Hepatocyte isolation

All experiments were undertaken in accordance with criteria outlined in a license granted under the Animals (Scientific Procedures) Act 1986, and approved by the Animal Ethics Committees of the University of Liverpool. Hepatocytes were isolated from male Wistar rats (Charles River, Manston UK) weighing approximately 250g. Rats were housed in controlled environmental conditions of 12 h light/dark cycle with free access to food and water. Rats were anesthetized with 200 μ L of pentoject by intraperitoneal injection and, a laparotomy performed. The hepatic portal vein was cannulated with a 20 g \times 48 mm vialon catheter (Becton Dickinson, U.K.), which was tied in place. Hanks balanced salt solution (HBSS) (Sigma, Poole, U.K.) without calcium and magnesium was perfused through the liver at 37 $^{\circ}$ C for 10 minutes during liver excision. Once excised, 100 mL of complete HBSS containing 50 mg of collagenase A (Roche, Welwyn Garden City, U.K.) was perfused through

the liver with recirculation until it was digested. The cells were dispersed into 100 mL Williams' E medium (WEM) (Sigma, Poole, U.K.) and filtered through a 125 μ m gauze into a sterile container. The cells were washed three times by centrifugation (100g for 2 minutes at 4 $^{\circ}$ C) and an aliquot was taken for counting. Cell suspensions (0.5×10^6 /mL) were prepared in WEM supplemented with 100 U/mL penicillin and 100 μ g/mL streptomycin, 2 mM/L glutamine (Sigma, Poole, U.K.), 10% fetal calf serum (Lonza biologics, Slough, U.K.), and 1 μ g/mL bovine insulin (Sigma, Poole, U.K.) and seeded at 70 000 cells/cm². The seeded plates were incubated at 37 $^{\circ}$ C in a humidified atmosphere containing 5% CO₂ for 3 h before the medium was exchanged for serum - and insulin-free WEM. Hepatocytes were cultured for 48 h as a monolayer on 100mm collagen I-coated culture dishes (Becton Dickinson, Oxford, UK). Medium was changed 3 h after seeding and each 24 h thereafter.

2.2 Sample protein preparation

For whole-cell preparations, 5×10^6 freshly-isolated hepatocytes or hepatocytes scraped into ice-cold phosphate-buffered saline (PBS) from two 100mm culture dishes, after 48h in culture, were washed twice with 5 mL PBS, by low speed centrifugation, 100g for 5 minutes at 4 $^{\circ}$ C. Cell pellets were sonicated directly in 200 μ L of 0.5M triethylammonium bicarbonate, 0.1% SDS (TEAB/SDS), with ten 1s pulses of a probe and the preparation clarified by centrifugation (10,000g, 2 minutes).

Nuclei were purified from freshly-isolated hepatocytes and cells after 48h culture. Nuclei were prepared from 5×10^7 freshly-isolated cells, or cells from at least ten 100mm culture plates. Cells were washed with ice-cold PBS, pellets were re-suspended in 5 mL ice-cold hypotonic solution (0.2x PBS) and placed on ice for 10 minutes. Hepatocytes were disrupted using a dounce homogeniser and were centrifuged at 1000g for 10 minutes at 4 $^{\circ}$ C. The pellet was re-homogenised in 5 mL 2M sucrose, 3.3 mM CaCl₂. 100 μ L of 2 M sucrose, 3.3mM CaCl₂ was pipetted into multiple micro-tubes and carefully overlaid with 400 μ L of sample homogenate and centrifuged at 18,000g, 4 $^{\circ}$ C for 1 h. The resulting pellets of nuclei were each washed with 0.5 mL PBS per tube, and centrifuged at 1000g for 10 minutes to form a pellet. The fraction was checked for the absence of whole cells and large debris by microscopic examination of a 5 μ L aliquot diluted 20-fold with PBS and a final concentration of 0.1% trypan blue. The pellets were combined and the proteins extracted by sonication in 30 μ L TEAB/SDS). The preparation was clarified by centrifugation (10,000g, 2 minutes). Protein concentration was determined in the supernatant using Bradford assay reagents (Sigma, Poole, UK).

2.3 iTRAQ labelling Cation exchange and Mass spectrometry

Labelling was carried out according to the manufacturer's instructions for 8-plex labelling (ABSciex, Warrington, UK)

using 75 µg cellular or nuclear protein per tag. For LC-MS/MS analysis of iTRAQ labelled samples, each cation exchange fraction was resuspended in 120 µL 5% ACN/0.05% trifluoroacetic acid (TFA) and 60 µL loaded on to the column. Samples were analysed on a QSTAR® Pulsar i hybrid mass spectrometer (ABSciex) and were delivered into the instrument by automated in-line liquid chromatography (integrated LC Packings System, 5 mm C18 nano-precolumn and 75 µm × 15 cm C18 PepMap column; Dionex, California, USA) via a nano-electrospray source head and 10 µm inner diameter PicoTip (New Objective, Massachusetts, USA). The pre-column was washed for 30 minutes at 30 µL/min with 5% ACN/0.05% TFA, prior to initiation of the solvent gradient in order to reduce the level of salt in the sample. A gradient from 5% ACN/0.05% TFA (v/v) to 60% ACN/0.05% TFA (v/v) in 70 min was applied at a flow rate of 300 nL/minute. The MS was operated in positive ion mode with survey scans of 1 second, and with an MS/MS accumulation time of 1 second for the three most intense ions. Collision energies were calculated on the fly based on the m/z of the target ion and the formula, collision energy = (slope × m/z) + intercept. The intercepts were increased by 3–5 V compared to standard data acquisition in order to improve the reporter ion intensities/quantitative reproducibility.

Data analysis was performed using ProteinPilot software (Version 3, ABSciex, Warrington, UK). The data were analysed with a fixed modification of MMTS-labelled cysteine, biological modifications allowed and with the confidence set to 10% to enable the false discovery rate to be calculated from screening the reversed SwissProt database 30-11-2009. Ratios for each iTRAQ label were obtained, using one sample of freshly-isolated cell or nuclear protein as the denominator in respective experiments. Raw iTRAQ files are available at the TRANCHE hash: yBAoOh3Q3eZ-CL8Dc10EJzm5m0vcxS27Lb4EUJYQB88/DGbpPpYkVVQxjvLmB74NomeFJR2CMvJX/jk6+gHne6lhaQAAAAAAAAAi7A==

2.4 Data processing and analysis

Proteins above the 1% false discovery rate, identified with ≥95% confidence (for those identified with more two or more peptides) or 99% confidence (for those identified by a single peptide) were included in the final data analysis. Only proteins meeting these criteria and identified in at least three samples were used in the statistical analysis in the R computational environment, version 2.14.1 [15]. The packages *marray* (16) and *multtest* [17] were used for comparison between freshly-isolated cells and cultured cells. R scripts are available at the TRANCHE hash yBAoOh3Q3eZCL8Dc10EJzm5m0vcxS27Lb4EUJYQB88/DGbpPpYkVVQxjvLmB74NomeFJR2CMvJX/jk6+gHne6lhaQAAAAAAAAAi7A==. Molecular functions were assigned using the PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system, utilising the gene expression analysis tool ‘compare

gene lists’ (www.pantherdb.org/) [18, 19].

2.5 Western Blotting

Proteins were denatured and reduced by heating to 80 °C for 10 minutes in Laemmli sample buffer (Sigma, Poole, U.K.) and 10 µg of each sample was resolved through a 10% polyacrylamide gel (Biorad, Hemel Hempstead, U.K.) and transferred to a nitrocellulose membrane (GE Healthcare, Slough, U.K.). After transfer, the nitrocellulose membrane was blocked overnight in Tris buffered saline, containing 0.1% Tween 20, (Sigma, Poole, U.K.) and 10% milk protein at 4 °C. Antibodies used were rabbit anti-Lamin A (Sigma, Poole, U.K.) and goat anti-HNF4 alpha (Insight biotechnology, Wembley, U.K.). Peroxidase-conjugated anti-rabbit, or anti-goat IgG antibodies (Sigma, Poole, U.K.) were used for chemiluminescence detection using ECL detection reagents and ECL hyperfilm (GE Healthcare, Slough, U.K.).

3. Results

3.1 The proteome of whole cell homogenates is enriched for molecular functions associated with liver.

iTRAQ analysis of freshly-isolated and cultured whole cell homogenates identified and quantified 431 proteins. The supporting data file, sheet ‘WC iTRAQ results’ details all proteins identified in whole cell extracts, their relative quantities, log₂ fold-change after culture and their associated p values (T test, comparing freshly-isolated and cultured samples). These proteins were compared against the *rattus norvegicus* reference list in PANTHER using the gene expression analysis tool ‘compare gene lists’. PANTHER’s background reference list of 27758 proteins mapped to 120 molecular functions and the experimental dataset mapped to 69 molecular functions (supporting data file, sheet ‘WC MF’). 41 molecular functions were represented by significantly more or fewer proteins ($p < 0.05$) than expected, based on the proportion of proteins from the reference list that mapped to each molecular function. Of these functions, 24 were significantly over-represented in the experimental dataset (Table 1) and 17 were under-represented. Many of the functions that are over-represented are functions associated with liver parenchyma, such as oxidoreductase-, peroxidase- and antioxidant activities. Under-represented molecular functions included those that are carried out by low abundance proteins, such as transcription regulation- or kinase activities.

3.2 Nuclear fractionation affects the number and percentage of identified proteins mapping to specific molecular functions.

The purity of hepatocyte nuclei isolation was checked by microscopic examination after addition of trypan blue to an aliquot of nuclei. Nuclei were seen as distinct organelles and the preparations were largely free from whole cell and cell

Table 1. Molecular functions significantly over-represented in the whole cell extract experimental dataset. The number of proteins assigned to each molecular function in the PANTHER classification system for the whole cell experimental dataset was compared to the number assigned in the reference rat proteome. Functions that were represented by significantly more proteins ($p < 0.05$) than expected based on dataset size are reported.

Molecular Function	REFLIST (27758)	WC dataset (419)	Expected	Observed/Expected	P-value
oxidoreductase activity	971	124	14.66	8.46	1.17E-76
catalytic activity	6555	248	98.95	2.51	1.94E-54
structural constituent of ribosome	814	54	12.29	4.39	1.97E-19
lyase activity	230	27	3.47	7.78	6.66E-16
transferase activity	2027	78	30.6	2.55	2.84E-14
isomerase activity	258	23	3.89	5.91	2.33E-11
racemase and epimerase activity	66	12	1	12.00	6.98E-10
hydro-lyase activity	71	12	1.07	11.21	1.57E-09
structural molecule activity	2389	72	36.06	2.00	1.57E-08
antioxidant activity	29	8	0.44	18.18	2.14E-08
peroxidase activity	27	7	0.41	17.07	2.48E-07
transaminase activity	27	7	0.41	17.07	2.48E-07
acyltransferase activity	209	15	3.15	4.76	1.05E-06
hydrogen ion transmembrane transporter activity	63	7	0.95	7.37	5.88E-05
transketolase activity	6	3	0.09	33.33	1.15E-04
protein disulfide isomerase activity	18	4	0.27	14.81	1.81E-04
transferase activity, transferring glycosyl groups	278	12	4.2	2.86	1.27E-03
ligase activity	652	20	9.84	2.03	2.57E-03
anion channel activity	62	5	0.94	5.32	2.73E-03
translation regulator activity	130	7	1.96	3.57	3.99E-03
translation factor activity, nucleic acid binding	133	7	2.01	3.48	4.51E-03
translation elongation factor activity	47	4	0.71	5.63	5.97E-03
transaldolase activity	1	1	0.02	50.00	1.50E-02
translation initiation factor activity	99	5	1.49	3.36	1.81E-02

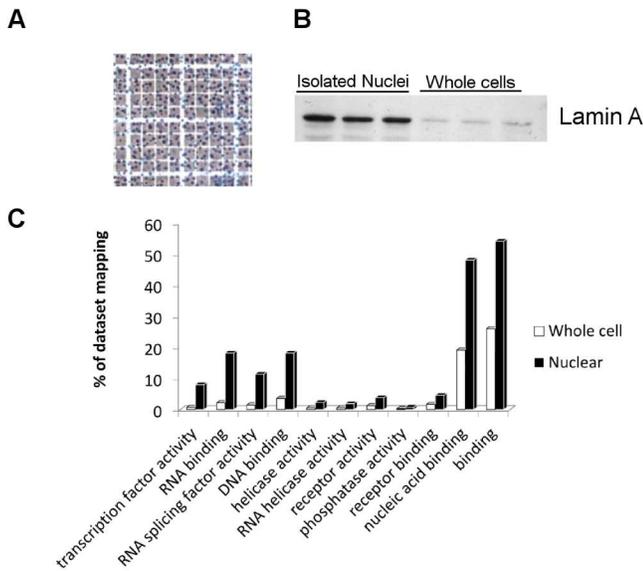


Figure 1. Enrichment of nuclei: A) Isolated nuclei were examined microscopically to verify purity and found to be largely free from whole cells and cell debris. B) Western blot for the nuclear envelope protein, Lamin A, demonstrates enrichment of nuclei. C) Bar chart of the molecular functions, identified in whole cell or nuclear preparations that were enriched 2-fold or more by nuclear fractionation. The number of proteins mapping to each function is expressed as a percentage of the total number of identified proteins in that experiment.

debris (figure 1A). As further confirmation of high levels of nuclear enrichment western blots were performed for the nuclear envelope protein, Lamin A. Immunodetection of Lamin A was much stronger in nuclear preparations compared with whole cell samples confirming the high degree of enrichment for nuclei (figure 1B). These highly pure preparations were used to analyse the hepatocyte nuclear proteome by iTRAQ proteomics. 268 distinct proteins were identified and quantified in our purified nuclear protein preparations. The supporting data file, sheet “Nuc iTRAQ results” details all proteins identified in nuclear preparations, their relative quantities, mean log₂ fold change in culture and associated *p* values. 156 proteins were unique to this experiment and not represented in the whole cell experiment. This dataset was also compared against the PANTHER rat reference list of proteins for mapping to molecular function (Supporting data file, sheet “Nuc MF”). There were 31 molecular functions represented by significantly more or fewer proteins (*p* < 0.05) than expected (23 molecular functions over-represented and 8 under represented). Table 2 details the molecular functions significantly over-represented by proteins in the nuclear dataset compared with expected values based on the reference list. Some of these molecular functions were over-or under-represented in common with the whole cell dataset but others are uniquely enriched or reduced in nuclear preparations. RNA binding, for example, was detected with the expected number of proteins in the whole cell dataset but more than 8-fold enriched in the nuclear preparations (table 2). Similarly transcription co-factor

activity, which was under-represented in whole cell samples, is more than 2-fold over-represented in the nuclear preparations relative to the reference dataset.

By simultaneously comparing the whole cell and nuclear datasets against the reference list of proteins in PANTHER we showed that nuclear fractionation changed the relative abundance of proteins associated with some molecular functions. This also gave an absolute count of the numbers of proteins that map to each function for each experiment. Many additional proteins were mapped to particular molecular functions depending on whether the protein preparation was whole cell- or nuclear-derived. In the whole cell dataset there were 37 molecular functions that were proportionally enriched compared to nuclear preparations. These functions were mainly associated with hepatocyte functions such as catalytic-, oxidoreductase- and transferase activities. Nuclear-derived protein preparations were proportionally enriched for 46 molecular functions (supporting data file, sheet “MF enrichment”). The functions identified in both datasets but proportionally enriched by more than 2 fold in nuclear samples are shown in figure 1 C. Many of the functions have a strong association with the nucleus, such as nucleic acid binding and transcriptional activity. Supporting data file, sheet “MF enrichment”, details, for all represented molecular functions in both datasets, the count for proteins mapping to that function for each dataset and the percentage of the each dataset mapping to that function.

3.3 Molecular function analysis of proteins changed in during culture.

Culture of hepatocytes caused a significant change in abundance for a large proportion of proteins in whole cell and nuclear samples. Using the quantitative data for whole cell extracts to construct a heatmap with unsupervised hierarchical clustering (figure 2A) demonstrated the clear differences between freshly-isolated and cultured samples with clustering of similar sample types and clear segregation by sample type. The mean log₂ fold-change for each protein was calculated and plotted as a frequency histogram (figure 2B). This showed that proteins were both up- and down-regulated with the larger changes occurring with lower frequency. Statistical analysis of these data, by T-test, showed that 206 proteins (>47% of the dataset) were significantly changed in abundance after 48h in culture (raw *p* < 0.05). After adjustment for multiple comparisons there were still 159 proteins (>36% of the dataset) significantly changed in abundance (BH *p* < 0.05). Log₂ fold-changes and *p* values are listed for each protein in the supporting data file, sheet “WC iTRAQ results” and summarised graphically as a volcano plot in figure 3A, and as a frequency histogram of adjusted *p* values in figure 3B.

The molecular functions of the significantly changed proteins in whole cell preparations were determined through PANTHER by selecting the complete list of identified proteins as a reference list. Of the 69 molecular functions repre-

Table 2. Molecular functions significantly over-represented in the nuclear extract experimental dataset. The number of proteins assigned to each molecular function in the PANTHER classification system for the nuclear extract experimental dataset was compared to the number assigned in the reference rat proteome. Functions that were represented by significantly more proteins ($p < 0.05$) than expected based on dataset size are reported.

Molecular Function	REFLIST (27758)	Nuclear dataset (264)	Expected	Observed / Expected	P-value
RNA binding	617	48	5.87	8.18	5.66E-29
nucleic acid binding	5086	127	48.37	2.63	3.60E-28
RNA splicing factor activity, transesterification mechanism	334	30	3.18	9.43	5.04E-20
catalytic activity	6555	120	62.34	1.92	6.65E-15
Binding	8663	143	82.39	1.74	9.89E-15
structural constituent of ribosome	814	30	7.74	3.88	3.84E-10
oxidoreductase activity	971	30	9.23	3.25	2.06E-08
antioxidant activity	29	5	0.28	17.86	1.02E-05
DNA binding	2682	48	25.51	1.88	1.53E-05
structural molecule activity	2389	43	22.72	1.89	4.03E-05
peroxidase activity	27	4	0.26	15.38	1.45E-04
translation factor activity, nucleic acid binding	133	7	1.26	5.56	3.26E-04
hydrogen ion transmembrane transporter activity	63	5	0.6	8.33	3.81E-04
translation initiation factor activity	99	6	0.94	6.38	4.18E-04
deaminase activity	38	4	0.36	11.11	5.23E-04
protein disulfide isomerase activity	18	3	0.17	17.65	7.29E-04
isomerase activity	258	9	2.45	3.67	9.33E-04
translation regulator activity	130	6	1.24	4.84	1.68E-03
RNA helicase activity	99	5	0.94	5.32	2.77E-03
transferase activity, transferring glycosyl groups	278	8	2.64	3.03	5.60E-03
helicase activity	170	6	1.62	3.70	6.15E-03
hydrolase activity, acting on ester bonds	780	13	7.42	1.75	3.75E-02
transcription cofactor activity	339	7	3.22	2.17	4.50E-02

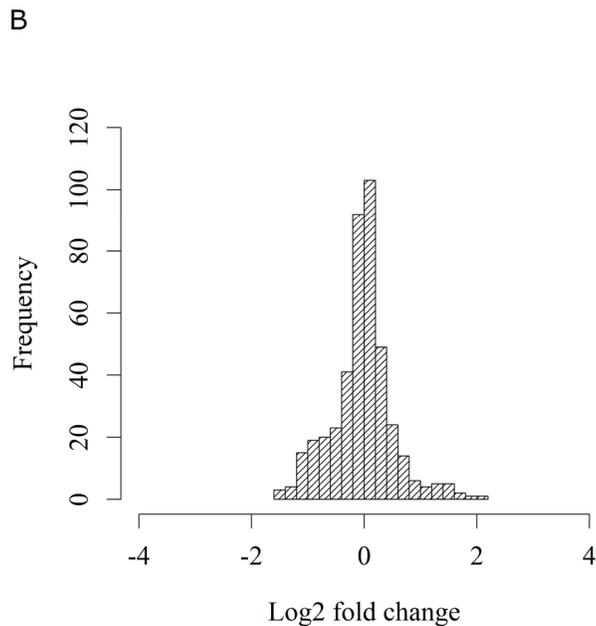
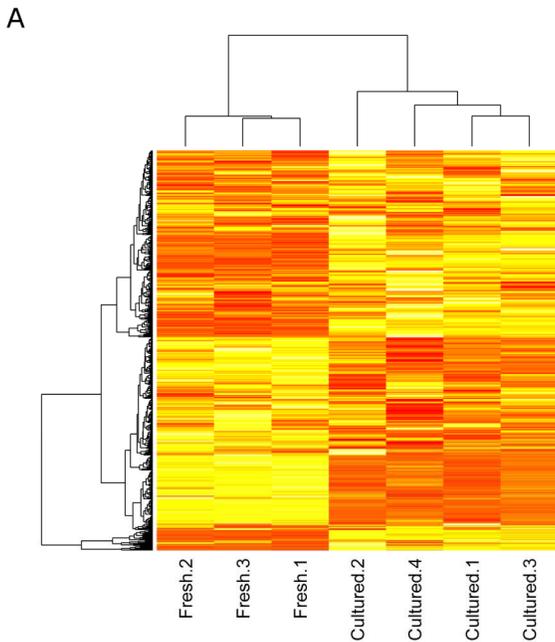


Figure 2. The changing proteome of cultured hepatocytes. Whole cell preparations were made from freshly-isolated hepatocytes or hepatocytes that were cultured for 48h. A) A heatmap with hierarchical clustering shows that samples segregate according to quantitative differences between freshly-isolated and cultured hepatocyte proteome. B) Frequency histogram for Log₂-fold change during hepatocyte culture. The larger changes occur with reduced frequency.

sented in the whole cell dataset, 65 functions were represented in the significantly changed subset when considering those with a raw *p* value <0.05, or 59 functions when using adjusted *p* values <0.05. 35 molecular functions were over-represented in this subset of proteins when compared against the complete whole cell dataset as a reference (supporting data file, sheet “WCPvalueMFs”).

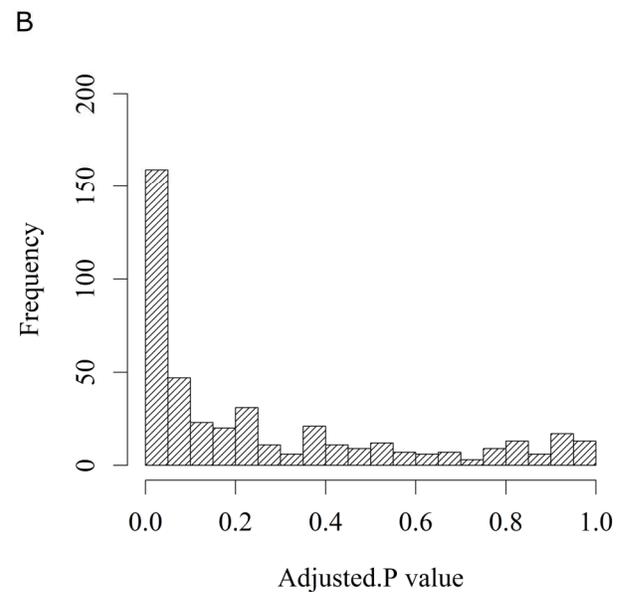
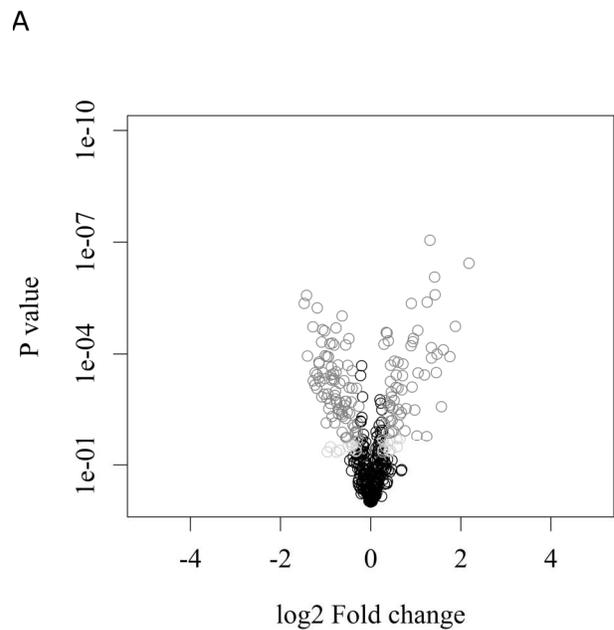


Figure 3. Statistical analysis of the changes in the cultured hepatocyte proteome. A) Volcano plot of Log₂ fold-change for each protein identified in whole cell preparations versus the significance (*p* value). Proteins not significantly changed are indicated with a black circle, those significantly changed +/- 20% with a raw *p* value <0.05 are coloured light grey and those significant after adjustment for multiple comparisons (Benjamini Hochberg *p* value <0.05) are dark grey. B) frequency histogram depicting the number of proteins that fall into each *p* value range.

3.4 Nuclear preparation extends the list of molecular functions that are identified as significantly changed in culture

Nuclear preparations also showed significant changes in a large proportion of proteins after 48h culture. Constructing a heatmap with unsupervised hierarchical clustering again demonstrated the clear differences between fresh and cul-

tured samples (figure 4A). The mean log₂ fold-change, plotted as a frequency histogram (figure 4B), shows that approximately equal numbers of proteins were up- and down-

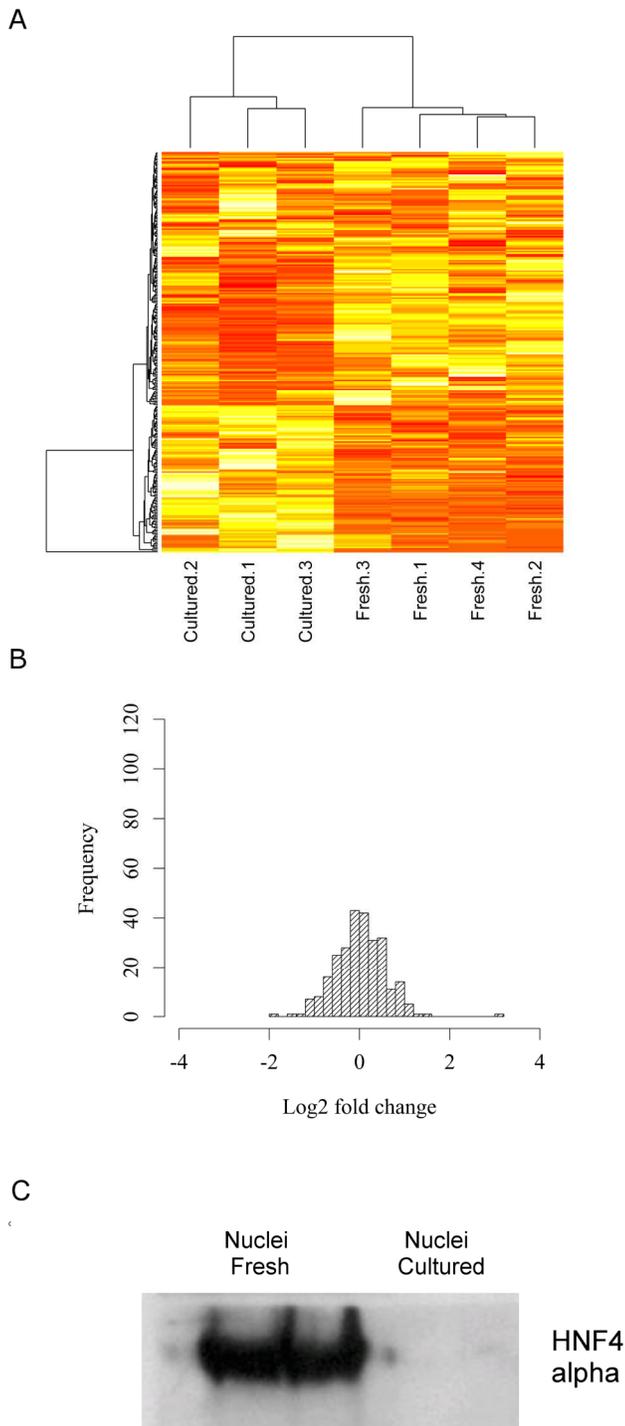


Figure 4. The changing nuclear proteome of cultured hepatocytes. A) Heatmap with hierarchical clustering shows the quantitative differences between freshly-isolated and cultured hepatocyte nuclear proteome. B) Frequency histogram for Log₂ fold change demonstrating the number of proteins up-regulated or down-regulated after culture. C) Western blot verification of reduced HNF4 alpha expression in cultured hepatocytes.

regulated. Hepatic nuclear factor 4 alpha was among the proteins identified was significantly reduced in expression in cultured cells. This was verified by western blot in 2 samples each of freshly-isolated and cultured nuclei (figure 4C). 99 proteins (>36% of the nuclear dataset) were significantly changed in abundance after 48h in culture (raw $p < 0.05$). After adjustment for multiple comparisons 58 proteins (>21%) remained significantly changed in abundance (BH $p < 0.05$). A total of 28 molecular functions were over-represented in the adjusted p value subset of data when compared against the complete nuclear dataset as a reference, including many functions performed in the nucleus (supporting data file, sheet “NucPvalueMFs”). Log₂ fold changes and p values are listed for each protein in supporting data file, sheet “Nuc iTRAQ results” and summarised graphically as a volcano plot in figure 5A, and as a frequency histogram of adjusted p values in figure 5B.

The functions of the changed proteins were assessed using PANTHER. The complete dataset from each experiment were used as a background list and the significantly changed proteins (raw $p < 0.05$ or BH $p < 0.05$) were compared for the numbers of proteins mapping to each molecular function. The results are summarised in figure 6 and the full results are given in supporting data file, sheet “nucPvalueMFs”.

4. Discussion

The short supply of primary human hepatocytes or alternative functional hepatocyte-like cells is a problem in drug development programs and is, in part, driving research into the development of stem cell-derived hepatocyte-like cells [20–24]. However, it is probable that any stem cell-derived model reaching hepatocyte-like status will suffer the same de-differentiation phenomenon as mature cells. It is therefore imperative to fully characterise the changes in primary cells in culture in order to identify possible interventions to prevent this loss of function.

Quantitative proteomic technologies such as iTRAQ [6, 25], isotope coded affinity tags [iCAT] [26] and stable isotope labelling by amino acids in cell culture (SILAC) [27] are becoming more and more accessible. The datasets typically encompass at least several hundred proteins, which represents the higher abundance or more easily detectable proteins. These multiplexed labelled proteomics technologies have the advantage of identifying and relatively quantifying proteins in multiple samples in a single experiment without gel casting and spot cutting associated with 2D gel-based proteomics. For our studies we use iTRAQ labelling as it allows relative quantification of hundreds of proteins in up to 8 samples simultaneously allowing statistical analysis of proteins from two experimental groups.

Many of the important proteins that control cellular phenotype and dynamic response are compartmentalised and/or are low abundance nuclear proteins. The nucleus consists of an estimated 14% of the entire cellular proteome [11] but many of its components are poorly represented in

datasets from un-fractionated cells. The method we employ results in clean nuclear preparations (Figure 1A). This reduced the amount of high abundance cytoplasmic/cytoskeletal proteins to allow the detection of more low abundance nuclear proteins. Others have performed experiments on purified nuclei [10-14] but to our knowledge this is

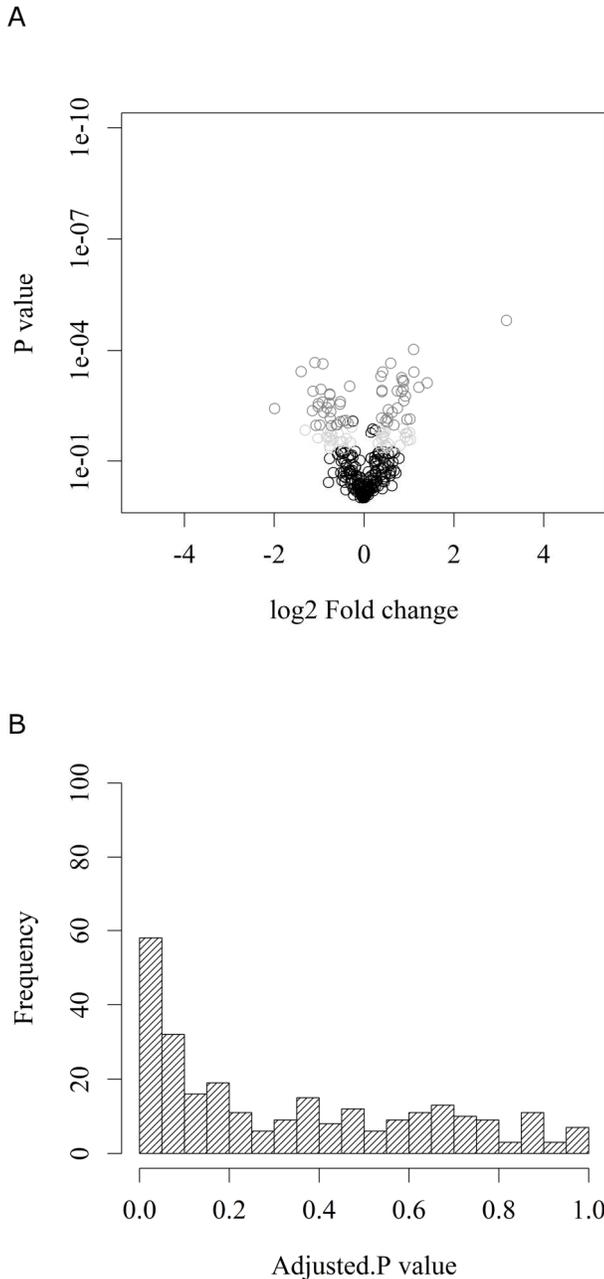


Figure 5. Statistical analysis of the changes in the cultured hepatocyte proteome: A) Volcano plot of Log₂ fold change for each protein identified in nuclear preparations versus the significance (*p* value). Proteins not significantly changed are indicated with a black circle, those significantly changed +/- 20% with a raw *p* value <0.05 are coloured light grey and those significant after adjustment for multiple comparisons (Benjamini Hochberg *p* value <0.05) are dark grey. B) frequency histogram depicting the number of proteins that fall into each *p* value range.

the first time primary hepatocyte nuclear preparations have been prepared for shotgun proteomics analysis to decipher the changes that occur to this organelle in hepatocytes during culture. This work therefore forms a basis for the understanding of the hepatocyte nuclear proteome and expands the knowledge of the hepatocyte proteins that change during culture, providing datasets describing the proteomic signature of hepatocyte de-differentiation.

We postulated that an altered nuclear proteome in cultured cells could underpin the change in the phenotype of primary hepatocytes during culture, as significant changes in the expression of hepatocyte-specific genes are evident after hepatocyte isolation and culture [6, 28-29]. Previous quantitative analysis of nuclear proteomes in other cell types has been undertaken but the scale of identification of changed proteins was smaller than presented here. Pewsey *et al* [12] used iTRAQ technology to investigate the changing nuclear proteome of differentiating embryonic carcinoma

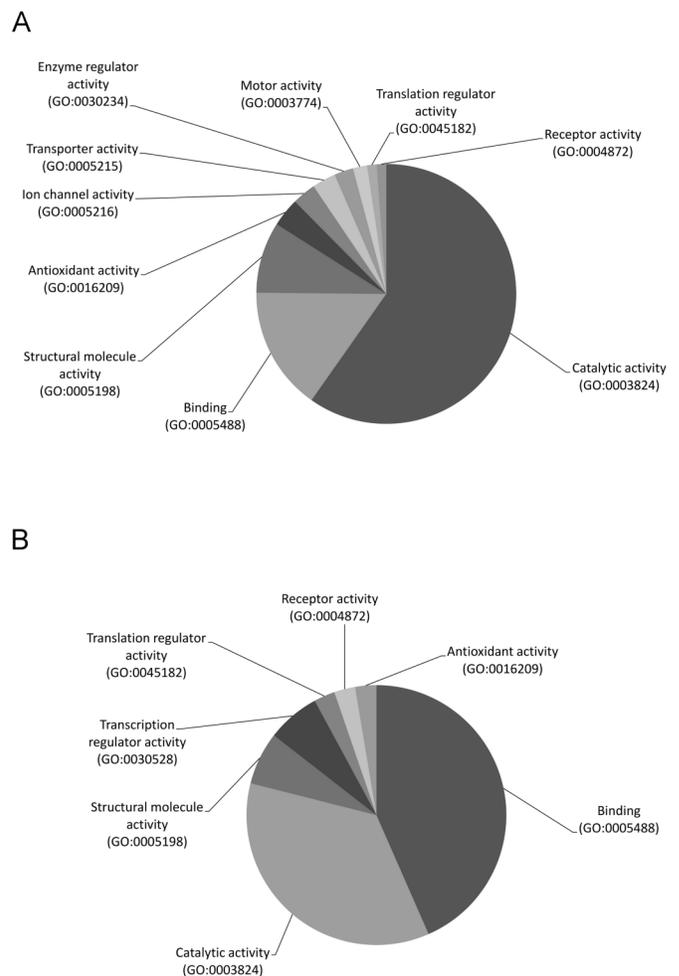


Figure 6. Molecular functions of proteins with alerted abundance in culture. Pie charts summarising the molecular functions that are performed by the proteins that were identified as significantly changed after culture (*p* value <0.05) in whole cell (A), or nuclear protein preparation (B). Significantly changed proteins were assigned to a molecular function using the Batch ID tool in PANTHER.

(NTERA-2) cells, identifying 37 proteins that were significantly altered in expression. Zhang *et al* [13] identified 22 differentially expressed nuclear proteins in cisplatin-treated HeLa cells by two-dimensional electrophoresis, using a similar approach Cao *et al* [14] identified 19 nuclear proteins changed in abundance by Epstein-Barr virus (EBV) infection. Besides our own work [6] shotgun proteomic study of changes in cultured primary hepatocytes is not extensively reported. Pan *et al* [30] performed a wide scale proteomics analysis of mouse primary hepatocytes with the purpose of comparing their phenotype with that of mouse hepatic cell lines but did not address the question of de-differentiation of the primary cells during culture.

Here we present proteomic datasets for freshly-isolated and cultured rat hepatocytes, generated using whole cells or purified nuclei. The value of this approach is clear with an expanded coverage of nuclear proteins and molecular functions represented by the 156 proteins that were uniquely detected in the nuclear fractions. The quantitative nature of these data allows a robust statistical analysis to identify quantitative changes that occur in the nucleus during cell culture. 64 proteins that were significantly changed in abundance after culture were identified in nuclei, but not detected in the whole cell experiment. These unique proteins were filtered for those annotated to the nucleus and the resulting 53 proteins are listed and their functions indicated in the supporting data file, sheet "UniqueSigNuc". These proteins include the transcription factors, established as vital to hepatic differentiation, HNF4alpha [9] and CCAAT/enhancer-binding protein beta [31], both of which were significantly reduced in cultured cells, as well as others with less well-defined roles in hepatocyte phenotype, but which are worthy of further investigation such as Nuclear factor 1 A-type and Methyl-CpG-binding protein 2 (supporting data file, sheet "Nuc iTRAQ results"). We mapped a greater number of proteins to molecular functions associated with the nucleus using the nuclear fractionation approach and by combining the protein lists for both experiments we increased the mapped molecular functions to a total of 83. However, proteins are redundant with respect to GO terms and an individual protein may appear in multiple categories. For example 'GO:0016563 : transcription factor activity' and 'GO:0030528: transcription regulator activity' were each represented by 3 proteins in the whole cell dataset and 21 proteins in the nuclear dataset, as the same proteins map to both terms. Nevertheless, this 7-fold increase in proteins detected in nuclear samples demonstrates the power of nuclear fractionation in improving the coverage of functions for a given organelle.

Sub-cellular expression proteomic profiling demands relatively large sample sizes in order to purify the required quantity of protein from organelles. This amount of protein is achievable with highly proliferative cells types, such as mammary epithelial cells [31], or proliferating

undifferentiated stem cells [32, 33], or proliferative immune cells [34]. Currently, the number of differentiated stem cells typically produced in a research laboratory-scale experiment prohibits such shotgun proteomic investigations. However, the 'top-down' approach we have applied here, to more fully characterise mature, differentiated cells is essential to establish the changes that occur as a result of de-differentiation in primary culture. Figures 6A and 6B summarise the molecular functions that are performed by the proteins that were significant changed with culture (BH $p < 0.05$), identified in each dataset by performing a Batch ID tool in PANTHER.

5. Concluding Remarks

Hepatocyte de-differentiation in culture is well documented [7] but still poorly understood. Despite great progress in hepatocyte culture systems over the past 20 years or more the conditions that allow the maintenance of physiologically relevant hepatocyte phenotype in culture for prolonged periods is still elusive. The application of global quantitative technologies is making it possible to undertake much broader phenotyping of cell models and assess what is changing during cell culture. Publication of datasets such as those presented here, will help address the question of what is changing in culture and aid the development of better hepatocyte model systems.

6. Supplementary material

A microsoft excel file is provided with supporting information on 9 sheets:

Sheet "WC iTRAQ results"; All proteins identified and quantified after applying filtering cut-offs as described in methods, for whole cell extract prepared from freshly-isolated and cultured hepatocytes. The values are reported are quantities relative to 1 fresh sample.

Sheet "WC MF"; Molecular functions performed by the proteins identified and quantified in whole cell (WC) hepatocyte experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function from the rattus norvegicus reference proteome in PANTHER. Over or under-representation indicated with a "+" or "-" respectively and the associated P value is reported.

Sheet "Nuc iTRAQ results"; All proteins identified and quantified after applying filtering cut-offs as described in methods, for nuclear protein extracts prepared from freshly-isolated and cultured hepatocytes. The values are reported are quantities relative to an experimental pool made from all nuclear samples.

Sheet "Nuc MF"; Molecular functions performed by the proteins identified and quantified in nuclear (Nuc) hepatocyte experiment. The number of proteins expected to fall into each category was estimated based on the proportion of

proteins mapping to each function from the *rattus norvegicus* reference proteome in PANTHER. Over or under-representation indicated with a "+" or "-" respectively and the associated *p* value is reported.

Sheet "MF enrichment"; The absolute count and percentage of identified proteins performing each Molecular function mapped by PANTHER.

Sheet "WCPvlaueMFs"; Molecular functions performed by the proteins identified as significantly changed in abundance after culture (T test raw *p* value or adjusted (BH) *p* value <0.05) in the whole cell proteomics experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function in the whole experimental dataset. Over or under-representation in the significantly changed sub-population is indicated with a "+" or "-" respectively and the associated *p* value is reported.

Sheet "NucPvalueMFs"; Molecular functions performed by the proteins identified as significantly changed in abundance after culture (T test raw *p* value or adjusted (BH) *p* value <0.05) in the nuclear proteomics experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function in the whole experimental dataset. Over or under-representation in the significantly changed sub-population is indicated with a "+" or "-" respectively and the associated *p* value is reported.

Sheet "UniqueSigNuc"; Proteins identified as significantly changed in abundance during culture were filtered for those annotated to the nucleus in the swissprot database. The change in abundance is given (log2), along with raw *p* value (t test), their annotated cellular locations and protein function, where known.

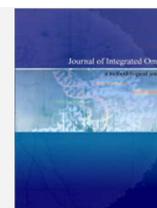
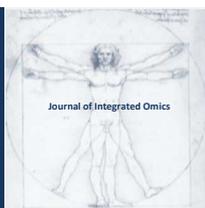
Acknowledgements

This work was funded by the MRC Centre for Drug Safety Science (grant number G0700654).

References

- Soars MG, McGinnity DF, Grime K, Riley RJ. *Chem Biol Interact* 2007;168:2-15.
- Mingoia RT, Nabb DL, Yang CH, Han X. *Toxicol In Vitro* 2007;21:165-173.
- Rogiers V, Vercruyse A. *Toxicology* 1993;82:193-208.
- LeCluyse EL. *Eur J Pharm Sci* 2001;13:343-368.
- Chesne C, Guyomard C, Fautrel A, Poullain MG, Fremond B, De Jong H, Guillouzo A. *Hepatology* 1993;18:406-414.
- Rowe C, Goldring CE, Kitteringham NR, Jenkins RE, Lane BS, Sanderson C, Elliott V, Platt, V. Metcalfe, P. Park, B. K. *Journal of proteome research* 2010;9:2658-2668.
- Elaut G, Henkens T, Papeleu P, Snykers S, Vinken M, Vanhaecke T, Rogiers V. *Current drug metabolism* 2006;7:629-660.
- Barthelery M, Salli U, Vrana KE. *Stem Cells Dev* 2007;16:905-919.
- Lemaigre F, Zaret KS. *Curr Opin Genet Dev* 2004;14:582-590.
- Dellaire G, Farrall R, Bickmore WA. *Nucleic Acids Res* 2003;31:328-330.
- Fink JL, Karunaratne S, Mittal A, Gardiner DM, Hamilton N, Mahony D, Kai C, Suzuki, H. Hayashizaki, Y. Teasdale, R. D. Towards defining the nuclear proteome. *Genome Biol* 2008;9:R15.
- Pewsey E, Bruce C, Tonge P, Evans C, Ow SY, Georgiou AS, Wright PC, Andrews, P. W. Fazeli, A. *Journal of Proteome Research* 2010;9:3412-3426.
- Zhang G, Sun L, Lu X, Chen Z, Duerksen-Hughes PJ, Hu H, Zhu X, Yang J. *Mutation research* 2012;746:66-77.
- Cao JY, Mansouri S, Frappier L. *Journal of virology* 2012;86:382-394.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini A J, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. *Genome biology* 2004;5:R80.
- Yang YH, Paquet A and Dudoit S. 2009. R package version 1.32.0;.
- Pollard KS, Gilbert HN. and Ge Y, Taylor S, and Dudoit S. R package version 2.10.0.
- Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff J A, Doremieux O. *Nucleic acids research* 2003;31:334-341.
- Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A, Lazareva-Ulitsky B. *Nucleic acids research* 2006;34:W645-650.
- Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, Gripon P, Kremsdorf D, Guguen-Guillouzo C, Corlu A. *Hepatology* 2007;45:957-967.
- Dalgetty DM, Medine CN, Iredale JP, Hay DC. *Am J Physiol Gastrointest Liver Physiol* 2009;297:G241-248.
- Hay DC, Fletcher J, Payne C, Terrace JD, Gallagher RC, Snoeys J, Black JR, Wojtacha D. Samuel K, Hannoun Z, Pryde A, Filippi C, Currie IS, Forbes SJ, Ross JA, Newsome PN, Iredale JP. *Proc Natl Acad Sci U S A* 2008;105:12301-12306.
- Basma H, Soto-Gutierrez A, Yannam GR, Liu L, Ito R, Yamamoto T, Ellis E, Carson SD, Sato S, Chen Y, Muirhead D, Navarro-Alvarez N, Wong RJ, Roy-Chowdhury J, Platt JL, Mercer DF, Miller JD, Strom SC, Kobayashi N, Fox IJ. *Gastroenterology* 2009;136:990-999.
- Baxter MA, Rowe C, Alder J, Harrison S, Hanley KP, Park BK, Kitteringham NR, Goldring CE, Hanley NA. *Stem Cell Res* 2010;5:4-22.
- Kitteringham NR, Abdullah A, Walsh J, Randle L, Jenkins RE, Sison R, Goldring CEP, Sanderson C, Williams S, Higgins L, Yamamoto M, Hayes J, Park BK. *J Proteomics* 2010.
- Jenkins RE, Kitteringham NR, Hunter CL, Webb S, Hunt TJ, Elsby R, Watson RB, Williams D, Pennington SR, Park BK. *Proteomics* 2006;6:1934-1947.
- Geiger T, Wehner A, Schaab C, Cox J, Mann M. *Molecular and Cellular Proteomics* 2012; 11(3):M111.014050. Epub 2012 Jan 25.
- Boess F, Kamber M, Romer S, Gasser R, Muller D, Albertini S, Suter L. *Toxicol Sci* 2003;73:386-402.
- Kienhuis AS, Wortelboer HM, Maas WJ, van Herwijnen M, Kleinjans JC, van Delft JH, Stierum RH. *Toxicol In Vitro*

- 2007;21:892-901.
30. Pan C, Kumar C, Bohl S, Klingmueller U, Mann, M. *Mol Cell Proteomics* 2009. **8**(3): 443-450.
31. Williams SC, Cantwell CA, Johnson PF. *Genes Dev* 1991;5:1553-1567.
32. Jaishankar A, Barthelery M, Freeman WM, Salli U, Ritty TM, Vrana KE. *Stem Cells Dev* 2009;18:793-802.
33. Sarkar P, Collier TS, Randall SM, Muddiman DC, Rao BM. *Proteomics* 2012;12:421-430.
34. Jarbouli MA, Wynne K, Elia G, Hall WW, Gautier VW. *Molecular immunology* 2011;49:441-452.



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.107

New insights in *Trypanosoma cruzi* proteomic map: further post-translational modifications and potential drug targets in Y strain epimastigotes

Daniela Gois Beghini¹, André Teixeira da Silva Ferreira¹, Vivian Corrêa de Almeida², Marcelle Almeida Caminha³, Floriano Paes Silva-Jr², Jonas Perales¹, Rubem Figueiredo Sadok Menna-Barreto^{*3}

¹Laboratório de Toxinologia, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil; ²Laboratório de Bioquímica de Proteínas e Peptídeos, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil; ³Laboratório de Biologia Celular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil.

Received: 24 August 2012 Accepted: 03 November 2012 Available Online: 09 November 2012

ABSTRACT

Chagas' disease is a neglected sickness endemic in Latin America, caused by the protozoa *Trypanosoma cruzi*. The current treatment for the disease is unsatisfactory, and the development of potent compounds for novel molecular targets is critical. In this framework, proteomics could be a powerful tool in the evaluation of possible candidates for drug intervention. In this work, a two-dimensional electrophoresis (2-DE) and mass spectrometry (MS) approaches were employed in *T. cruzi* epimastigotes (Y strain). Different gel staining protocols (Coomassie Blue, Pro-Q-Diamond and Pro-Q-Emerald) were performed to assess the protein content and possible post-translational modifications of this parasite. Here, 78 most intense spots were identified by Coomassie staining, 22 by Pro-Q-Diamond (phosphoproteins) and 15 by Pro-Q-Emerald (glycoproteins). Compared with the results of other large-scale *T. cruzi* proteomic studies, 15 novel proteins were identified here using MALDI-TOF/TOF, and 12 of these have not yet been described at the protein level. Functional analysis of the identified proteins pointed to protein metabolism, and the localisation prediction indicated cytosol as the most prevalent localisation of these proteins. Eight proteins presented no similarity to human sequences and thus represent a group of promising biomolecules for chemotherapy intervention. Our data provides novel insights in the metabolic pathways of *T. cruzi*, which could aid in the discovery of alternative drugs for Chagas' disease.

Keywords: *Trypanosoma cruzi*; Chagas' disease; Chemotherapy; Post-translational modifications; Mass spectrometry; Proteomics .

1. Introduction

Chagas' disease is a neglected tropical disease that is endemic in Latin America and was first reported in an ancient Chinchorro civilisation 9,000 years ago [1]. This illness is caused by infection with its etiological agent, the hemoflagellate protozoan *Trypanosoma cruzi*. The *T. cruzi* life cycle involves invertebrate and vertebrate hosts as well as the three evolutive forms of the protozoa. In the triatomine midgut, epimastigotes proliferate and subsequently migrate to the

posterior intestine, where differentiation into the infective metacyclic trypomastigote form occurs. After haematophagy, the infective metacyclic forms reach the bloodstream of the vertebrate host where they can infect the cells of various mammalian tissues. In the intracellular environment, the parasites differentiate into proliferative amastigotes. Subsequently, another differentiation occurs to produce trypomastigotes, which are responsible for disseminating the in-

*Corresponding author: Dr. R.F.S. Menna-Barreto, Laboratório de Biologia Celular, Instituto Oswaldo Cruz, FIOCRUZ, Av. Brasil 4365 – 21040-360, Manguinhos, Rio de Janeiro, Brazil; Phone number: 00 55 21 25621393; Fax number: 00 55 21 25621432; E-mail address: rubemb@ioc.fiocruz.br

fection. The transmission of Chagas' disease primarily occurs by infection with the triatomine vector, but it also has been reported to result from blood transfusion, congenital transmission, organ transplantation, the ingestion of contaminated food and laboratory accidents [2]. Due to its anthropic action, the triatomine favours a domiciliary and peridomiciliary behaviour, increasing human exposure and risk of infection. A recent report stated that 8 million people in Latin America have this disease and described the major concern that immigration-related globalisation will bring infected people to non-endemic regions, such as Europe, North America and Australia [2].

The disease treatment is still based on benznidazole and nifurtimox, nitroderivatives that were empirically developed forty years ago and present important limitations, such as high doses, severe side effects and controversial efficacy in chronic patients [3]. Intense efforts have been directed to discover new compounds against *T. cruzi*, better characterise their particular mechanisms of action and identify the molecules and pathways involved [4].

Gene expression in trypanosomatids, including *T. cruzi*, is very unusual in that the open reading frames are arranged in long polycistronic arrays and monocistronic mRNAs are created by post-transcriptional processing. Thus, the regulation of gene expression in *T. cruzi* is mainly post-transcriptional [5]. Previous studies have revealed nontranslated mRNA in the *T. cruzi* cytoplasm [6]. Therefore, the use of nucleic acid-based tools, such as RNA microarrays and PCR, to study gene expression is limited in trypanosomatids, and thus, the application of proteomic techniques is highly desirable.

Proteomics is a high-throughput approach that is extensively employed in structural and functional studies on the regulation of protein expression to validate genome annotations [7]. Different stage-specific molecules and proteins associated with parasite drug resistance have been identified by proteomic approaches in *T. cruzi* [8-11], reinforcing the importance of these techniques in the selection of metabolic targets of new inhibitors or therapeutic drugs [2,12].

In the present study, we used two-dimensional electrophoresis (2-DE) and mass spectrometry to contribute additional information to the descriptive proteomic map of *T. cruzi* epimastigotes (Y strain). The Y strain is a natural strain of *T. cruzi* I that is found in Brazil and presents high parasitaemia and mortality in mouse models. Despite its virulence, the Y strain is susceptible to benznidazole treatment [13]. Here, *T. cruzi* samples analysed by 2-DE were stained with different dyes to reveal specific phosphoproteins and glycoproteins, which were then identified by MS. The identified *T. cruzi* glycoproteins are particularly important as targets for therapeutic intervention, and our analysis of the phosphorylated proteins could provide crucial insights into the signalling networks that govern the metabolism of this parasite. The identified peptides and proteins were characterised using bioinformatic tools in an attempt to find potential drug targets.

2. Materials and methods

2.1 Parasites and sample preparation

T. cruzi epimastigotes (Y strain) were maintained in liver infusion tryptose (LIT) medium at 28°C and harvested during exponential growth phase. The parasites were harvested by centrifugation, washed three times with phosphate buffered saline (PBS, pH 7.4) and incubated in sample lysis solution (7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris, 60 mM DTT, 1% ampholytes) containing the Complete Mini protease inhibitor cocktail (Roche Applied Science, Indianapolis, USA). The parasite cells were further disrupted via ten cycles of freezing in liquid nitrogen and thawing at room temperature, with gentle mixing between cycles, as previously described [11]. To obtain the soluble protein fraction, the cell lysate was centrifuged at 13,000 g for 30 min, and the protein content was determined using the 2D Quant kit (GE Healthcare, Buckinghamshire, England).

2.2 Two-dimensional electrophoresis and image analysis

Isoelectric focusing of the soluble protein fraction from the epimastigote extracts was performed using an IPGphor system (GE Healthcare) and 18-cm IPG strips of pH 4-7. The protein fraction was diluted in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 60 mM DTT, 1% ampholytes and a trace amount of bromophenol blue) and loaded onto each strip. The strips were rehydrated for 12 h at 30 V and isoelectrically focused using the following protocol: 200 V/1 h, 500 V/1 h, 1000 V/1 h, a gradient of 1000 to 8000 V for 10 min and 8000 V/6 h. The strips were then equilibrated in 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% v/v glycerol, 2% w/v SDS, 0.002% w/v bromophenol blue and 1% w/v DTT for 15 min, followed by a second 15-min incubation in the same buffer but with 4% iodoacetamide instead of DTT. Separation in the second dimension was performed by SDS-PAGE using 12% acrylamide gels at 2.5 W/gel for 30 min and then 100 W until completion using a DALTsix system (GE HealthCare). The proteins separated on the 2D gels were stained to detect phosphoproteins, glycoproteins or total protein as described below.

2.3 Detection of total protein

Gels were fixed three times with 30% ethanol and 2% phosphoric acid for 30 min per wash, followed by three successive washes in 2% phosphoric acid in water for 20 min. Then, the gels were incubated in 18% ethanol, 2% phosphoric acid and 15% ammonium sulphate for 30 min, followed by the addition of 2% Coomassie Brilliant Blue G in water (Sigma-Aldrich, St. Louis, USA). The gels were scanned and the spots were detected, quantified and analysed with Image Master Platinum software (GE Healthcare). Each spot was manually inspected and, when necessary, edited with the Edit Spot Tools to improve accuracy.

2.4 Detection of phosphoproteins

2D gels were fixed in 50% methanol and 10% acetic acid overnight, washed three times with deionised water for 10 min per wash and incubated in Pro-Q Diamond phosphoprotein stain (Molecular Probes) for 90 min. To reduce the background signal, gels were destained with three successive washes of 20% acetonitrile and 50 mM sodium acetate (pH 4.0). Images were acquired with a Typhoon Trio system (GE Healthcare) using a 532 nm laser for excitation and a 580 nm filter for emission. Following image acquisition, the gels were stained with colloidal Coomassie as previously described for the detection of total protein to view and excise the protein spots that were revealed by this specific staining.

2.5 Detection of glycoproteins

Pro-Q Emerald 488 (Molecular Probes) was used to detect glycoproteins in the 2D gels. Briefly, gels were fixed with two changes of 50% methanol and 5% acetic acid for 1 hour per wash, followed by incubation in 1% periodic acid and 3% acetic acid for 1 hour to oxidise the glycans on these proteins. After washing in 3% acetic acid, gels were incubated in a Pro-Q Emerald dye solution for 2 hours in the dark and washed again in 3% acetic acid to reduce the gel background. All steps were carried out at room temperature with gentle agitation, and the images were captured on a Typhoon Trio system using a 488 nm laser for excitation and a 520 nm filter for emission. Following image acquisition, as described above for Pro-Q Diamond staining, the gels were stained with colloidal Coomassie for the visualisation and excision of the protein spots revealed by the glycoprotein-specific staining.

2.6 In-gel tryptic digestion, sample desalting and MALDI-TOF TOF spectrometry analysis

Coomassie-, Pro-Q-Emerald- and Pro-Q-Diamond-stained spots were excised and digested as previously described with some modifications [14]. Protein spots were excised from the gel and washed with a 1:1 (v/v) 50 mM ammonium bicarbonate (pH 8.0)/acetonitrile solution, followed by shaking for 15 min. This washing procedure was repeated with fresh solution until destaining was complete, and the gel was dehydrated by the addition of 200 μ L of acetonitrile for 5 min. Each sample was then rehydrated with approximately 10 μ L of ice-cold trypsin solution (20 ng/ μ L in 40 mM ammonium bicarbonate [pH 8.0]) and left on ice for 1 hour. The incubation was allowed to proceed for 16 h at 37 $^{\circ}$ C. After the digestion, the peptides were extracted twice by the addition of 30 μ L of 50% acetonitrile and 5% formic acid and transferred to 0.6 mL tubes. Each sample was concentrated in a vacuum centrifuge to a final volume of 5-10 μ L. C18 ZipTip micropipette tips (Millipore, Bedford, USA) were used to desalt the peptides. The tips were first activated with acetonitrile and then equilibrated with 0.1% trifluoro-

acetic acid (TFA) in water. The samples were aspirated and dispensed eight times, and the tips were washed 3-5 times with 0.1% TFA in water. The peptides retained in the tips were eluted with 1.5 μ L of 50% acetonitrile plus 0.1% (v/v) TFA in water. Each eluate was immediately spotted on the ABI 192-target MALDI plate (Applied Biosystems, USA) by co-crystallisation with 0.3 μ L of the α -cyano-4-hydroxycinnamic acid matrix (CHCA) (10 mg/mL CHCA in 50% acetonitrile, 0.3 [v/v] TFA in water).

Raw data for protein identification were obtained using the 4700 Proteomics Analyzer (Applied Biosystems, Foster City, CA). Both MS and MS/MS data were acquired in positive and reflectron mode using a neodymium-doped yttrium aluminium garnet (Nd:YAG) laser with a 200-Hz repetition rate. Typically, 1,600 shots were accumulated for spectra in MS mode, while 3,000 shots were accumulated for spectra in MS/MS mode. Up to ten of the most intense ion signals with a signal to noise ratio above 20 were selected as the precursors for MS/MS. External calibration in MS mode was performed using a mixture of four peptides: des-Arg1-Bradykinin (m/z = 904.47), angiotensin I (m/z = 1,296.69), Glu1-fibrinopeptide B (m/z = 1,570.68) and ACTH (18-39) (m/z = 2,465.20). MS/MS spectra were externally calibrated using known fragment ion masses observed in the MS/MS spectrum of angiotensin I. A search of the MS/MS database was performed against the NCBI databases using Mascot software (www.matrixscience.com). The search parameters included the allowance of two missed tryptic cleavages and non-fixed modifications of methionine, tryptophan, histidine (oxidation) and cysteine (carbamidomethylation and propionamide). The peptide per sample plate (pps) and peptide per well (ppw) files were generated from the raw (or native) MS data according to the following parameters using Data Explorer Software (Applied Biosystems). For MS1: mass range, 900-4,000 Da; peak density, 15 peaks per 200 Da; signal-to-noise ratio, 20; minimum area, 100 μ m²; maximum peaks per spot, 60. For MS2: mass range, 60 or 20 Da for the precursor; peak density, 55 peaks per 200 Da; signal-to-noise ratio, 2; minimum area, 10; maximum peaks per precursor, 200. The protein identifications based on the MS/MS data were validated using Scaffold 2 software (Proteome Software Inc., Portland, OR). The identified proteins were accepted if they possessed >90.0% probability, as specified by the Peptide Prophet algorithm [15]. Protein probabilities were assigned by the Protein Prophet algorithm [16].

2.7 Bioinformatics analysis

A BLAST search of the genome project databases was performed for *Leishmania major* (<http://www.genedb.org/Homepage/Lmajor>), *Trypanosoma brucei brucei* (<http://www.genedb.org/Homepage/Tbruceibrucei427>) and *Homo sapiens* (<http://www.ncbi.nlm.nih.gov/genome/>) to find annotated sequences to all *T. cruzi* polypeptides identified in this study. Alternatively, to suggest a possible function of the identified hypothetical *T. cruzi* proteins, another BLAST

analysis of the whole NCBI database was performed. The function and localisation of the polypeptides was predicted using Gene Ontology (GO, <http://www.geneontology.org/>), the Kyoto Encyclopaedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>), Uniprot (<http://www.uniprot.org/>) and SwissProt (<http://www.expasy.org/tools/>). The high hit sequences (E value $\leq 10^{-50}$) of *L. major*, *T. brucei* and *H. sapiens* were aligned to the identified *T. cruzi* proteins using the multiple sequence alignment (MSA) program T-Coffee (European Bioinformatics Institute, Cambridge, UK). Different sequences (≥ 20 amino acids) observed between the *T. cruzi* and *H. sapiens* proteins were classified as non-homologous. Motif prediction analysis was performed using query sequences in NCBI Conserved Domain

(<http://www.ncbi.nlm.nih.gov/cdd>) and ScanProsite tools (<http://www.expasy.ch/tools/scanprosite/>).

3. Results

The protein preparation and two-dimensional electrophoresis conditions for *T. cruzi* epimastigote samples have been previously established (Menna-Barreto et al., 2010), allowing for high reproducibility and effective spot identification with excellent ion scores. In the colloidal Coomassie-stained gel (pH 4-7), 600 spots were detected and 117 most intensely staining spots were identified as 78 distinct proteins (Figure 1). The complete list of the identified proteins is shown in supplementary Table S1. Between these identified proteins,

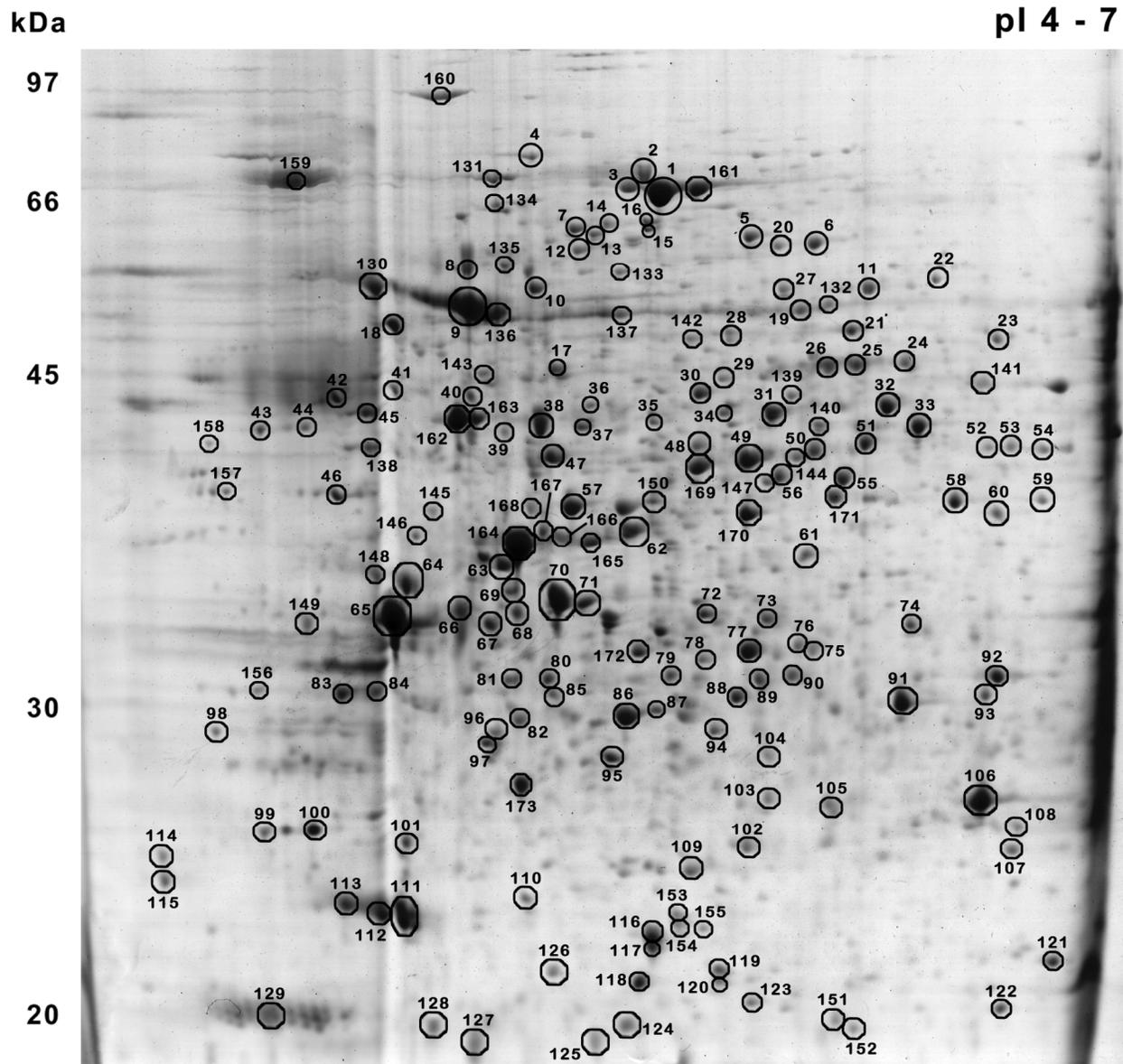


Figure 1. Colloidal Coomassie Blue staining of *T. cruzi* epimastigotes (Y strain). 2-DE was performed with 500 μ g of soluble protein using pH 4-7 IPG strips (18 cm), followed by 12% SDS-PAGE. Circles and numbers represent proteins and refer to the spot identification in Table S1 (supplementary data).

fifteen novel polypeptides not previously addressed in any *T. cruzi* proteome investigation were showed. These proteins are: mitochondrial heat shock, hypothetical proteins 1-6, glycosyl transferase family 2, metallocoxy-peptidase chain A, protein of unknown function DUF909, unnamed protein product, chaperonin GroEL, twin-arginine translocation pathway signal serine carboxypeptidase, conserved hypothetical protein and Sorting nexin GRD19 homolog. Interestingly, among these molecules, only a few had been described at the protein (3/15) or transcriptional (1/15) level. Most of these proteins were predicted (10/15) or even inferred by homology (1/15); this is the first description of their expression in this protozoan.

Here, we showed a map of the post-translational modifications revealed by staining 2-D gels with two different fluorescent dyes for phosphoproteins and glycoproteins. Three hundred and eighty-four phosphoproteins (Pro-Q-Diamond) and 34 glycoproteins spots (Pro-Q-Emerald) were detected. However, only 34 (Pro-Q-Diamond) and 24 (Pro-Q-Emerald) spots could be linked to their corresponding protein, resulting in 22 and 15 different proteins, respectively (Figures 2A,B). The complete list of identified proteins is presented in supplementary Table S2. To further investigate the two most common PTMs, phosphorylation and glycosylation, 2D gels were stained with Pro-Q-Diamond and Pro-Q-Emerald, respectively, leading to 31 different identifications, as shown in Table S2. A general evaluation of the previous status of the phosphoproteins and glycoproteins identified here demonstrated a similar percent distribution of proteins described as predicted (11/31), inferred by homology (8/31)

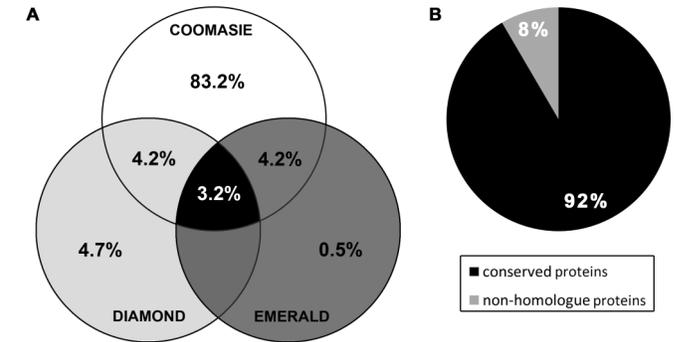
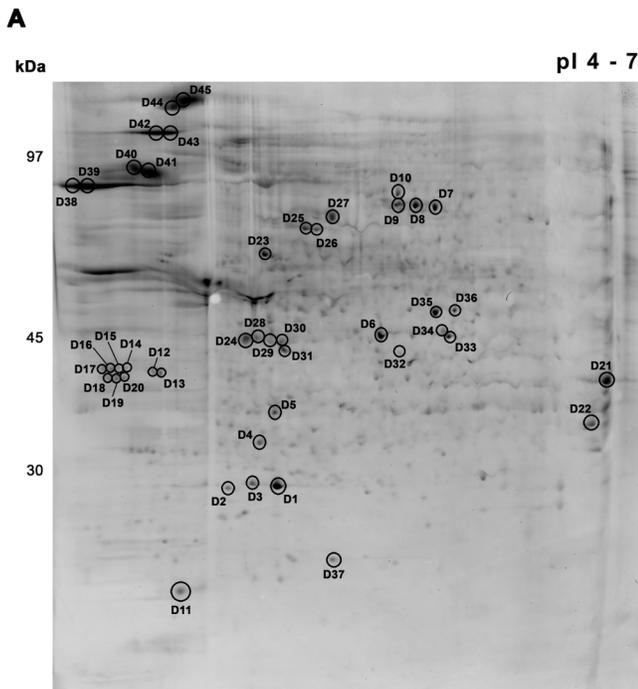


Figure 3. Percentage of protein identification in *T. cruzi* epimastigote (Y strain) after staining with colloidal Coomassie Blue, Pro-Q-Diamond and Pro-Q-Emerald.

or even detected at the protein (5/31) or transcriptional (5/31) level, reinforcing the necessity and relevance of further high throughput studies on post-translational modifications in protozoan pathogens, including *T. cruzi*. Surprisingly, some proteins were only identified as phosphoproteins (4.7%) and glycoproteins (0.5%) after the use of a specific fluorescent dye (Figure 3A), which could be explained by an increase in colloidal Coomassie sensitivity when the staining was performed after the labelling of the samples with fluorescent dyes (data not shown).

The percentage of peptides identified through each of the three staining protocols varied: colloidal Coomassie staining led to 83.2% of the peptide identifications, followed by Pro-Q-Diamond staining (8.9%) and Pro-Q-Emerald staining (4.7%). Only 3.2% of spots were identified by all of the three protocols (Figure 3). A mascot search of the whole NCBI

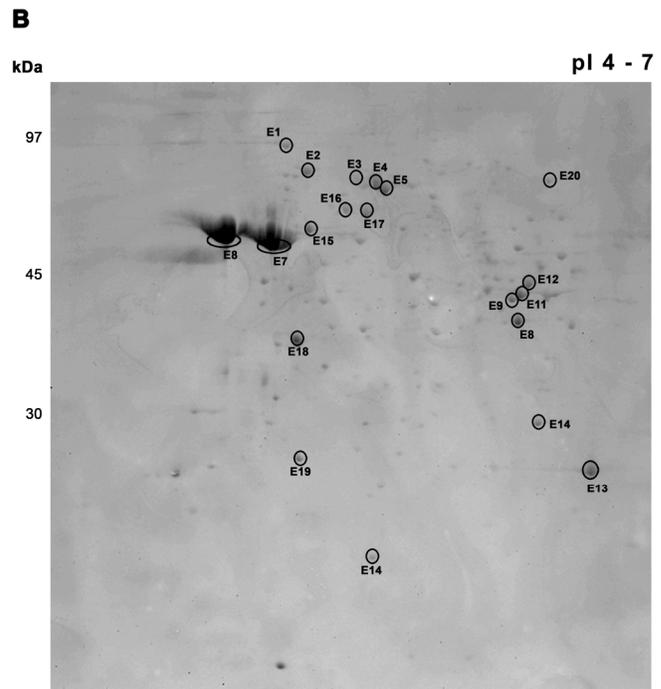


Figure 2. Pro-Q-Diamond (A) and Pro-Q-Emerald (B) staining of *T. cruzi* epimastigotes (Y strain). 2-DE was performed with 500 µg of soluble protein using pH 4-7 IPG strips (18 cm), followed by 12% SDS-PAGE. Circles and numbers represent proteins and refer to the spot identification in Table S2 (supplementary data).

database evidenced that 82.2% of the identified proteins were specific to *T. cruzi*, being a low percentage of identifications (6.7%) corresponded to highly conserved proteins that were not associated with any specific organism (data not shown). Compared to the *L. major*, *T. brucei* and *H. sapiens* genome databases, a high percentage of proteins were conserved in trypanosomatids; 92% of the identified proteins shared high sequence identity or similarity to *H. sapiens*. A low percentage of polypeptides (8%) did not appear to be homologous to any human peptide sequence (Figure 3B). After the MSA analysis, 21 candidates were selected as potential drug targets by their low similarity to *H. sapiens* proteins. Among those proteins, 10 presented no identity to humans, which two were excluded as potential targets, and another 11 proteins presented promising sequence fragment with no identity to human proteins (Table 1). The motif prediction analysis showed that these sequence fragments represent sites of post-translational modifications, such as phosphorylation, amidation, N-glycosylation and/or N-myristoylation.

Among all of the *T. cruzi* identifications, 10 hypothetical proteins were found and numbered from 1 to 10 (Figures 1 and 2A, Tables S1 and S2). A BLAST analysis of the whole NCBI database suggested the possible predicted protein functions, as shown in Table S3. Localisation and functional analysis demonstrated that cytosolic proteins were the most abundant (41%), followed by proteins localised in the mitochondrion (15.4%), nucleus (9.0%), glycosome (5.1%), plasma membrane (3.8%), endoplasmic reticulum (2.6%), flagellum (2.6%), reservosome (1.3%), cytoskeleton (1.3%) and proteasome (1.3%) under our experimental conditions (Figure 4A). In relation to predicted function, the most common pathway was the protein metabolism pathway, represented by the protein biosynthesis pathway (26.9%) and degradation pathway (19.2%), equalling 46.1%

of the total sequences. The other proteins were grouped into pathways including energetic metabolism (23.1%), motility and intracellular trafficking (9%), redox balance (2.6%), cell signalling (2.6%), lipid metabolism (1.3%), nucleic acid regulation (2.6%) and the polyamine pathway (2.6%) (Figure 4B).

4. Discussion

More than one century after the discovery of Chagas' disease, this illness still represents a serious health problem not only for its endemic region but also for developed countries due to the immigration of infected people. The current treatment of the disease is unsatisfactory, and an efficient prophylactic therapy is required, reinforcing the importance of designing new drug candidates for treatment and identifying alternative drug targets that are specific for the parasite [3,4,17]. In this work, we provide additional information on the proteomic map of *T. cruzi* epimastigotes through the use of fluorescent dyes and bioinformatics analysis.

Over the last decade, the proteomic map of different strains of *T. cruzi* epimastigotes has been extensively investigated, including clone Dm28c and the sylvan isolates [7-10, 18-23]. About Y strain epimastigotes, the unique large-scale protein investigation was performed in 2008 in order to characterise the molecules involved in the resistance and susceptibility of these parasites to the current chemotherapeutic agent benznidazole. Fifty-five proteins were identified, with approximately 10% hypothetical proteins [10], similar percentage to that observed in the present work. Here, also employing the 2-D gels followed by MALDI-TOF/TOF analysis, it was identified 78 distinct proteins approximately 30% more proteins.

Among the 10 trypanosomatid-specific proteins, each one represents a good candidate for drug intervention due to

Table 1. Motif analysis of *T. cruzi* sequences

PROTEIN	MOTIF FUNCTION ^a
heat shock protein 70	N-myristoylation, phosphorylation
thiol transferase Tc52	phosphorylation
ATPase beta subunit	N-myristoylation, phosphorylation
serine carboxypeptidase	phosphorylation
unnamed protein product	N-myristoylation, phosphorylation
arginine kinase	N-glycosylation, phosphorylation
cystathionine beta-synthase 6	amidation, phosphorylation
dihydrolipoamide acetyltransferase	phosphorylation
aminopeptidase	N-myristoylation, phosphorylation
arginase superfamily protein	N-myristoylation, phosphorylation
RNA-binding protein RGGm	N-myristoylation, phosphorylation, amidation

^a Sequences of >20 amino acids that are absent in human homologues.

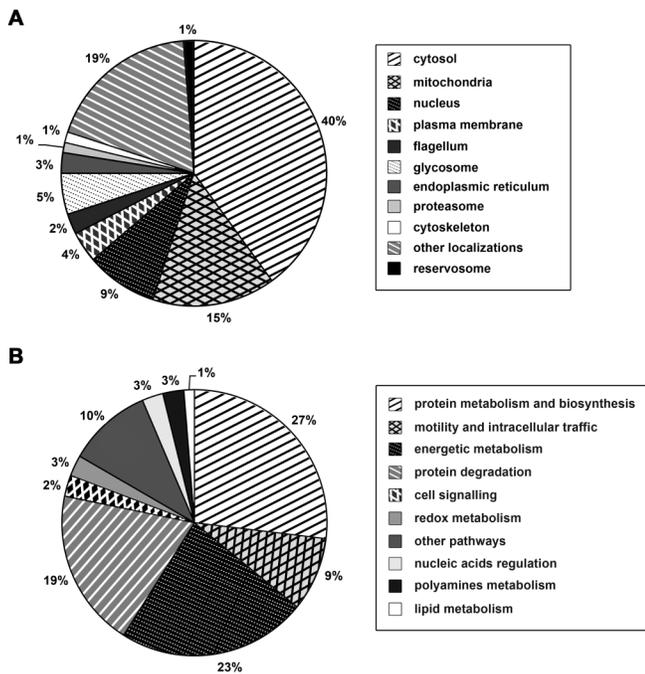


Figure 4. Localisation and functional classification of *T. cruzi* epimastigote (Y strain) proteins. (A) Organellar localisation of all proteins identified in this study. (B) Known or predicted function of each identified protein.

their inhibition could lead to the parasite death by a different pathway. Trypanothione synthetase is a crucial molecule in the main anti-oxidant pathway in trypanosomatids [24], related to ROS balance. 24-C-methyltransferase participates in ergosterol biosynthesis [25], directly involved in the parasite's plasma membrane fluidity. Calpain is a protease associated to the protozoa differentiation and infectivity [26,27]. A predicted dehydrogenase was better characterised in 2006 as prostaglandin F2 alpha synthase [28], being related to detoxification processes. Phosphoenolpyruvate carboxykinase is a key enzyme in gluconeogenesis [29]. Nucleoside phosphorylase participates in purine metabolism, being fundamental for nucleoside recovery after nucleotide degradation [30,31]. HslVu is a subunit of a large complex that degrades misfolded proteins called the proteasome [32]. I/6 autoantigen is a calcium-binding microtubule-associated protein responsible for crosslinking microtubule filaments and involved mitosis, vesicle trafficking and other events [33].

The last two candidates are involved in bioenergetics and for different reasons are not promising targets. P28 is a predicted pyruvate kinase superfamily protein that presents almost a half of the identical sequence of pyruvate kinase. The high similarity of pyruvate kinase to the corresponding mammalian protein led us to exclude P28 as a potential molecular target for novel trypanocidal drugs. Cytochrome c oxidase is a fundamental enzyme of respiratory electron transport chain [34]. Despite the differences between human and *T. cruzi* complex IV, the presence of alternative oxidases in the parasite [35], suggests that this protein is not a good target for drug design.

Comparison of the identified *T. cruzi* epimastigote proteins with other trypanosomatid and human sequences allowed for the selection of 21 proteins as potential drug targets, among which 10 displayed no hits in a BLAST analysis against human database and these trypanosomatid-specific proteins will further discussed here. The other 11 proteins presented stretches of more than 20 amino acids that did not align to the respective human sequences in the MSA, but the possibilities of misalignment due to the presence of multimeric proteins indicated that better algorithms must be used in order to increase the confidence.

PTMs of proteins are well characterised to be functionally important for various physiological and pathological processes, such as signal transduction and cell-cell recognition [36]. Until this study, Y strain epimastigotes' glycoprotein content has not been assessed. In 2006, the glycoproteome of trypanosomatids (Brazil strain) was published [37], demonstrating the presence of 31 glycosylated proteins, absent here. Because glycosylation could be fundamental for host cell recognition of the protozoa, it is expected that the infective parasite form presented a different pool of molecules in the relation to the insect form. Nakayasu and co-workers (2009) preliminarily assessed phosphoproteome of Y strain epimastigotes by LC-MS/MS [38], identifying 107 proteins that was not found here. Surprisingly, 19 of the phosphoproteins identified here by fluorescent staining of 2D gels and MALDI-TOF/TOF were not in the previous Y strain identification list. These molecules are heat shock-like 85, alpha tubulin, peroxiredoxin, paraflagellar rod protein 1D, glutamine synthetase, actin, pyrroline-5-carboxylate synthetase, 14-3-3 protein, chain A of spermidine synthase, enolase 1, eukaryotic initiation factor 5a, chaperonin containing T-complex protein, 80 kDa prolyl oligopeptidase, aminoacylase, proteasome beta 3 subunit and hypothetical proteins 7, 8, 9 and 10.

5. Concluding remarks

Despite many previous efforts to describe the proteome of epimastigotes, new molecules are still being identified in this proliferative form of *T. cruzi*. Our data could provide new insights into the signalling networks of this protozoan, supplying additional information about its cell biology. Here, it was identified 8 promising good candidates for drug intervention, being necessary a validation in the mammalian forms of parasite.

6. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/107/0>.

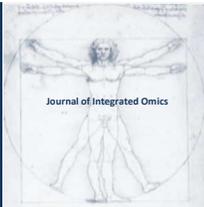
Supplementary Material includes Tables S1, S2, S3 and MS/MS spectra data.

Acknowledgments

We are very thankful to Nadja Oliveira and Thabata Duque for their help in the comparative proteome analysis. This work was supported with grants from CNPq, FAPERJ, PDTIS/FIOCRUZ and FIOCRUZ.

References

1. A.C. Aufderheide, W. Salo, M. Madden, J. Streitz, J. Buikstra, F. Guhl, B. Arriaza, C. Renier, L.E. Jr Wittmers, G. Fornaciari, M. Allison, PNAS 101 (2004) 2034-2039.
2. S.L. De Castro, D.G. Batista, M.M. Batista, W. Batista, A. Daliry, E.M. de Souza, R.F.S. Menna-Barreto, G.M. Oliveira, K. Salomão, C.F. Silva, P.B. Silva, M.N. Soeiro, Mol. Biol. International (2011), 306928.
3. J.R. Coura, S.L. De Castro, Mem. Inst. Oswaldo Cruz 97 (2002) 3-24.
4. M.N. Soeiro, S.L. De Castro, Exp. Opin. Ther. Targets 13 (2009) 105-121.
5. C. Clayton, M. Shapira, Mol. Biochem. Parasitol. 156 (2007) 93-101.
6. F.B. Holetz, L.R. Alves, C.M. Probst, B. Dallagiovanna, F.K. Marchini, P. Manque, G. Buck, M.A. Krieger, A. Correa, S. Goldenberg, FEBS J. 277 (2010) 3415-3426.
7. M. Ferella, D. Nilsson, H. Darban, C. Rodrigues, E.J. Bontempi, R. Docampo, B. Andersson, Proteomics 8 (2008) 2735-2749.
8. A. Parodi-Talice, R. Durán, N. Arrambide, V. Prieto, M.D. Piñeyro, O. Pritsch, A. Cayota, C. Cerveñansky, C. Robello, Int. J. Parasitol. 34 (2004) 881-886.
9. J.A. 3rd Atwood, D.B. Weatherly, T.A. Minning, B. Bundy, C. Cavola, F.R. Opperdoes, R. Orlando, R.L. Tarleton, Science 309 (2005) 473-476.
10. H.M. Andrade, S.M. Murta, A. Chapeaurouge, J. Perales, P. Nirdé, A.J. Romanha, J. Proteom. Res. 7 (2008) 2357-2367.
11. R.F. Menna-Barreto, D.G. Beghini, A.T. Ferreira, A.V. Pinto, S.L. De Castro, J. Perales, J. Proteomics 73 (2010) 2306-2315.
12. J.A. Urbina, Acta Tropica 115 (2010) 55-68.
13. R. Devera, O. Fernandes, J.R. Coura, Mem. Inst. Oswaldo Cruz 98 (2003) 1-12.
14. A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, M. Mann, PNAS 93 (1996) 14440-14445.
15. A. Keller, A.I. Nesvizhskii, E. Kolker, R. Aebersold, Analyt. Chem. 74 (2002) 5383-5392.
16. A.I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, Analyt. Chem. 75 (2003) 4646-4658.
17. J.C.P. Dias, Cadernos de Saúde Pública 23 (2007) 13-22.
18. J. Paba, J.M. Santana, A.R. Teixeira, W. Fontes, M.V. Sousa, C.A. Ricart, Proteomics 4 (2004) 1052-1059.
19. E.M. Cordero, E.S. Nakayasu, L.G. Gentil, N. Yoshida, I.C. Almeida, J.F. da Silveira, J. Proteom. Res. 8 (2009) 3642-3652.
20. C.L. Sodr , A.D. Chapeaurouge, D.E. Kalume, L. Mendonça-Lima, J. Perales, O. Fernandes, Arch. Microbiol. 191 (2009) 177-184.
21. S.A. Kikuchi, C.L. Sodr , D.E. Kalume, C.G. Elias, A.L. Santos, M.N. Soeiro, M. Meuser, A. Chapeaurouge, J. Perales, O. Fernandes, Exp. Parasitol. 126 (2010), 540-551.
22. E.S. Nakayasu, T.J. Sobreira, R.Jr Torres, L. Ganiko, P.S. Oliveira, A.F. Marques, I.C. Almeida, J. Proteome Res. 11 (2011) 237-246.
23. D. P rez-Morales, H. Lanz-Mendoza, G. Hurtado, R. Mart nez-Espinosa, B. Espinoza, J. Biomed. Biotech. (2012) 902803.
24. F. Irigo n, L. Cibils, M.A. Comini, S.R. Wilkinson, L. Floh , R. Radi, Free Rad. Biol. Med. 45 (2008) 733-742.
25. J.A. Urbina, International Symposium on the Centennial of the Discovery of Chagas' Disease. Rio de Janeiro, 2009.
26. V. Ennes-Vidal, R.F. Menna-Barreto, A.L. Santos, M.H. Branquinha, C.M. d'Avila-Levy, J. Antimicrob. Chemother. 65 (2010) 1395-1398.
27. V. Ennes-Vidal, R.F. Menna-Barreto, A.L. Santos, M.H. Branquinha, C.M. d'Avila-Levy, PLoS One 6 (2011) e18371.
28. S.M. Murta, M.A. Krieger, L.R. Montenegro, F.F. Campos, C.M. Probst, A.R. Avila, N.H. Muto, R.C. Oliveira, L.R. Nunes, P. Nird , O. Bruna-Romero, S. Goldenberg, A.J. Romanha, Mol. Biochem. Parasitol. 146 (2006), 151-162.
29. S. Trapani, J. Linss, S. Goldenberg, H. Fischer, A.F. Craievich, G. Oliva, J. Mol. Biol. 313 (2001) 1059-1072.
30. R.L. Miller, C.L. Sabourin, T.A. Krenitsky, Biochem. Pharmacol. 36 (1987) 553-560.
31. R.G. Silva, M.J. Veticcatt, E.F. Merino, M.B. Cassera, V.L. Schramm, J. Am. Chem. Soc. 133 (2011) 9923-9931.
32. C. Gille, A. Goedel, C. Schloetelburg, R. Preif ner, P.M. Kloetzell, U.B. Gobel, C. Frommell, J. Mol. Biol. 326 (2003) 1437-1448.
33. E. Detmer, A. Hemphill, N. M ller, T. Seebeck, Eur. J. Cell Biol. 72 (1997) 378-384.
34. P. Brzezinski, R.B. Gennis, J. Bioenerg. Biomemb. 40 (2008) 521-531.
35. J.J. Wen, N.J. Garg, Antioxid. Red. Sign. 12 (2010) 27-37.
36. M. Mann, O.N. Jensen, Nature Biotech. 21 (2003) 255-261.
37. J.A. 3rd Atwood, T. Minning, F. Ludolf, A. Nuccio, D.B. Weatherly, G. Alvarez-Manilla, R. Tarleton, R. Orlando, J. Proteome Res. 5 (2006) 3376-3384.
38. E.S. Nakayasu, M.R. Gaynor, T.J. Sobreira, J.A. Ross, I.C. Almeida, Proteomics 9 (2009) 3489-3506.



JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.108

Normalization of protein at different stages in SILAC subcellular proteomics affects functional analysis

Kerry L. Inder¹, Dorothy Loo¹, Yu Zi Zheng², Robert G. Parton³, Leonard J. Foster², Michelle M. Hill^{1*}

¹The University of Queensland Diamantina Institute, Brisbane, Queensland, 4102, Australia; ²Centre for High-Throughput Biology and Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4; ³Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia.

Received: 30 August 2012 Accepted: 12 November 2012 Available Online: 16 November 2012

ABSTRACT

Quantitative subcellular proteomics is a powerful method to interrogate spatial dynamics of cells or tissues. Stable isotope labeling by amino acids in cell culture (SILAC) is a popular quantitative approach that is ideally suited to subcellular proteomics because samples can be combined very early to reduce technical variability in the subcellular fractionation and downstream processing. However, validation of results using orthogonal methods such as immunoblotting do not allow mixing of samples prior to fractionation, leading to potentially different outcomes. Here we have investigated the impact protein normalization before or after subcellular fractionation has on the functional analysis and experimental conclusions. As a model system, we compared the detergent-resistant membrane (DRM) fraction of mouse embryonic fibroblasts (MEF) from caveolin-1-null mice with wildtype controls. Caveolin-1 is cholesterol-binding protein which is essential for formation of plasma membrane caveolae, a subtype of lipid raft membrane microdomains. Surprisingly, we found that the relative protein content of DRM as a percentage of total protein content is 1.6 fold higher for Cav1^{-/-} MEF compared to wild type MEF, leading to different SILAC ratios in pre fractionation mix and post fractionation mix experiments. Most of the observed differences were replicated by mathematical modeling of the normalization effect, with the striking exception for mitochondrial DRM proteins. Interestingly, caveolin-1 affected DRM proteins in the post fractionation mix data showed a significant enrichment of the mitochondrial oxidative phosphorylation pathway, which was not observed in the pre fractionation mix experiment. The observed quantitative changes in mitochondrial DRM proteins using different analyses suggest a caveolin-1 induced change rather than simple contamination, and may support recent reports of caveolin-1-dependent mitochondrial cholesterol changes. Based on these results, we recommend a thorough understanding of how experimental conditions impact relative subcellular fraction in order to make an informed decision on the most appropriate point to combine SILAC samples for quantitative subcellular proteomic analysis.

Keywords: Quantitative subcellular proteomics; Organellar proteomics; Caveolin-1; Caveolae; Lipid raft; Detergent resistant membranes.

Abbreviations:

SILAC, Stable isotope labeling by amino acids in cell culture; **DRM**, Detergent resistant membrane; **MS**, Mass spectrometry; **MEFs**, Mouse embryonic fibroblasts.

1. Introduction

The functional organization of the cell is central to its biological activity. Trafficking to specific subcellular compartments facilitates the functional activity of proteins. While

some proteins predominantly localize to a particular organelle, many proteins traffic to multiple locations where they exert distinct functions [1,2]. Understanding the spatial con-

*Corresponding author: Michelle M. Hill, Level4, R-wing, Princess Alexandra Hospital Ipswich Rd, Woolloongabba, QLD, 4102; E-mail address: m.hill2@uq.edu.au; Phone number: +61 7 3176 7774; Fax number: +61 7 3176 5946.

figuration of proteins and how this influences cell processes has therefore been an important question since the advent of cell biology. Over the last decade, proteomics has emerged as a powerful method to study the total protein complement of cells or tissue. More recently, proteomics has proven to be a successful approach to characterizing the repertoire of proteins within subcellular compartments [3,4]. This technology presents researchers with a high-throughput method to assign protein localizations. Furthermore, a particular organelle of interest can be focused upon or instead changes in protein distribution after perturbations can be assessed.

In general, subcellular proteomics is achieved through subcellular fractionation to enrich for organelle/s of interest. Well established protocols are available for enrichment of most subcellular compartments. Organelles such as the nucleus and mitochondria are easily purified based on their density and size and as such the protein complement of these organelles has been well documented using proteomics [5-7]. However, enrichment of some organelles such as the endoplasmic reticulum (ER) and Golgi apparatus are less straight-forward resulting in considerable levels of contaminating proteins [8]. Often, validation using techniques such as western blotting and electron microscopy is needed to distinguish bona fide organelle components from contaminating proteins. Subcellular fractionation also reduces sample complexity and increases detection of low abundant proteins. Whole cell analyses contain large amounts of protein which generates an overwhelming number of peptides. This is further complicated by large deviation in protein expression levels, ranging up to as much as 10 orders of magnitude [9]. As a result, low abundance proteins, which are often important regulatory proteins such as kinases, are concealed by highly abundant proteins. To date, organellar proteomics has been performed on the phagosome, lysosome, lipid rafts, exosomes, plasma membrane, clathrin-coated vesicles, spliceosome, nuclear pore, nucleolus, ER-Golgi intermediate compartment (ERGIC) and peroxisome, highlighting the depth and scope of this technique [3,8].

Recently, quantitative approaches have been applied to subcellular proteomics enabling the comparison of organelles under different experimental conditions. Stable isotope labeling by amino acids in cell culture (SILAC) is particularly well suited for subcellular proteomics. This method involves metabolic labeling of cells in culture with different isotopic labeled amino acids that can be distinguished by the mass spectrometer [10]. One advantage of SILAC is that samples can be mixed prior to the subcellular fractionation processes, eliminating variation between sample handling. The reduction in sample number simplifies and speeds up the otherwise time consuming and often cumbersome subcellular fractionation protocols. In addition, the disparity between mass spectrometry runs is also circumvented, significantly reducing the error rate.

Combining quantitative methods with subcellular proteomics has allowed the spatial dynamics within cells to be interrogated on a large scale. Emmott *et al.* [11] used SILAC

to study the host subcellular proteome in response to infection with Coronavirus Infectious Bronchitis Virus. Work by Dhungana and colleagues [12] focused on the detergent resistant membrane (DRM), a fraction enriched in lipid rafts, of macrophages in response to lipopolysaccharide treatment. They found that compartmentalization and activation of the 26S proteasome in DRM mediates activation of the MAPK pathway. More recently, we used quantitative SILAC and subcellular proteomics to investigate the role of caveolin-1, an integral membrane protein, in the aggressive prostate cancer cell line PC-3 [13]. We analyzed total membrane, DRM, prostasome, and secreted fractions. Our results suggested a role for caveolin-1 in modulating the lipid raft environment that accentuates secretion pathways possibly via ER sorting.

In the current study, we have investigated the widely accepted practice of mixing SILAC samples prior to subcellular fractionation and importantly, the impact of this on the functional analysis and experimental conclusions. In our model we have studied the protein caveolin-1 which is a major structural protein of caveolae. Caveolae are specialized lipid raft microdomains on the plasma membrane that are characterized by their flask shaped invaginations [14]. Caveolae are involved in many cellular processes including endocytosis, lipid regulation, and signal transduction [14]. Caveolin-1 deficient mouse embryonic fibroblasts (MEFs) have been widely used to ascertain the functions of caveolin-1 and caveolae [15, 16]. Here we have employed SILAC and subcellular proteomics to compare wild type MEFs and caveolin-1 deficient MEFs. We focused on DRM that includes caveolae and non caveolae lipid rafts and investigated the consequence of mixing SILAC samples before or after fractionation. Interestingly our results clearly demonstrate dramatic differences in the functional analysis between the different mixing steps. Our results caution the general practice of mixing SILAC samples prior to fractionation and instead recommend a thorough understanding of the changes in biology caused by experimental treatments in order to make an informed decision on the most appropriate point to combine SILAC samples for quantitative subcellular proteomic analysis.

2. Material and Methods

2.1 Cell culture and SILAC

Immortalized wild type (WT) and Caveolin-1-null (Cav1-/-) MEFs were generated as previously described [16]. MEFs were grown and maintained in DMEM lacking Lysine and Arginine (Sigma) with 10% dialyzed FBS (Bovogen) and supplemented with the following amino acids: '0/0' for the normal isotopic Lys and Arg (Sigma) and '8/10' for $^{13}\text{C}_6^{15}\text{N}_2$ -Lys and $^{13}\text{C}_6^{15}\text{N}_4$ -Arg (Silantes). Cell populations were amplified 200-fold in the labeling media to achieve > 99% incorporation as confirmed by LC-MS/MS. For each analysis, two 15 cm plates of cells were used for detergent-resistant mem-

brane extractions.

2.2 Detergent-resistant membrane (DRM) preparation

Detergent resistant membranes (DRMs) were prepared as previously described [13].

2.3 LC-MS/MS, database searching and SILAC quantitation

Protein samples (30ug) were separated on a 10% SDS-PAGE to 8-9 mm and stained with colloidal coomassie. Protein gel slices (1 mm) were excised and de-stained with a solution of 50% acetonitrile in 25mM ammonium bicarbonate before reduction with 20mM DTT (Sigma) and alkylation with 50mM IAA (Sigma). The gel pieces were subsequently pH adjusted with 50mM ammonium bicarbonate and dehydrated before overnight digest at 37°C with 0.01mM Trypsin (Promega) in a buffer of 10% acetonitrile and 50mM ammonium bicarbonate. Samples were extracted with 60% acetonitrile, 1% formic acid, dried in a speed-vac and resuspended in 10µl of 5% v/v formic acid for LC-MS/MS.

Peptides were analyzed using a 1200 Series nano HPLC and Chip-Cube Q-TOF 6510 (Agilent Technologies). Peptides were resolved using the Agilent large capacity HPLC chip (G4240-62010) 150 mm 300 Å C18 chip with 160nL trapping column. A 45 minute gradient from 10% to 45% solvent B was used. Solvent A was 0.1% formic acid and solvent B was 0.1% formic acid and 90% acetonitrile. HPLC loading pump was set to 3% B, flow rate of 4uL/min while analytical pump was set to 10%B and flow rate of 0.3uL/min. Q-TOF mass spectrometer was programmed to acquire 8 MS spectra/sec and 4MS/MS spectra/sec with dynamic exclusion after 2 MS/MS and released after 0.2min.

Mass spectrometry data was analyzed using Spectrum Mill (Agilent, B.04.00.127) search engine. Data was extracted with carbamidomethylation cysteine and SILAC amino acids N-Lys, $^{13}\text{C}_6^{15}\text{N}_2$ -Lys, N-Arg and $^{13}\text{C}_6^{15}\text{N}_4$ -Arg as a fixed/mix modifications. Extracted data were searched against SwissProt (release-2010_03 containing 23,000 entries) mouse database with carbamidomethylation cysteine and SILAC amino acids N-Lys, $^{13}\text{C}_6^{15}\text{N}_2$ -Lys, N-Arg and $^{13}\text{C}_6^{15}\text{N}_4$ -Arg as a fixed/mix modifications as appropriate, and oxidized methionine as variable modifications. Precursor and product mass tolerance was set to +/- 20ppm and \pm 50ppm respectively. Reverse database scores were calculated with Spectrum Mill search engine. All peptides identified had a global false discovery rate of less than 0.9%. Protein identification cut-offs were set to protein score > 11, peptide score > 10 and scored peak intensity > 60%.

Single peptide identifications were excluded from further analysis. Mean SILAC ratio (L/H) and standard deviation was calculated using all the peptide ratios matched to a protein, and p-values were calculated using the peptide SILAC ratios.

2.4 Assignment of gene ontology and functional enrichment analysis

Proteins with a SILAC p-value < 0.05 were submitted to GeneGo for identification of Gene Ontology (GO) terms over represented in each list. Correction for multiple hypothesis testing was performed by controlling for the False Discovery Rate at p = 0.05. Subcellular localization was assigned using UniProt.

2.5 Mathematical modeling

Representative SILAC fold changes (20 values in total) ranging from -10, -8, -6, -4, and -2 fold decrease to a 2, 4, 6, 8, and 10 fold increase were used in the model. Half of the SILAC fold changes (10 values) were a SILAC fold change of 1 since the majority of proteins remain unchanged. The pre fractionation values were adjusted by 1.6 fold to account for the increase contribution from Cav1-/- MEFs. The post fractionation values were left unchanged. The predicted outcomes were plotted as a line graph.

3. Results and Discussion

3.1 Sample preparation and mass spectrometry analysis

Using SILAC subcellular proteomics, we investigated if normalization of protein amount before and after subcellular fractionation would impact upon the functional analysis. To this end, we compared the subcellular proteomes of WT MEFs and Cav1-/- MEFs. Given that caveolin-1 is necessary for caveolae formation, we analyzed DRM, a fraction enriched in all cholesterol dependent lipid raft domains. Typical of SILAC subcellular proteomics, we combined equal amount of protein after cell lysis but prior to fractionation from SILAC-labeled WT MEFs and Cav1-/- MEFs, then iso-

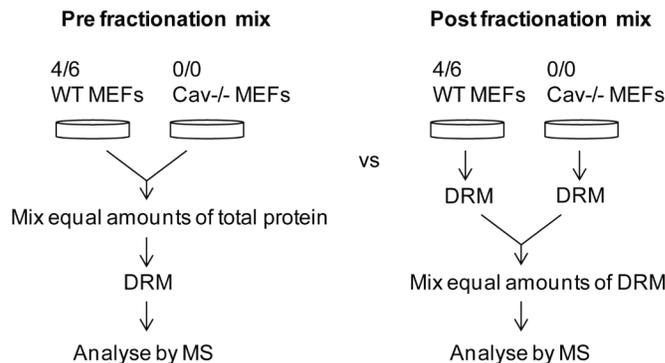


Figure 1. Work flow comparing mixing SILAC samples before or after subcellular fractionation. Typical of most SILAC subcellular proteomics experiments, SILAC labeled cells were lysed and equal levels of total protein were combined pre fractionation. DRM were purified from the combined sample and then analyzed by mass spectrometry (MS). In comparison, DRM was isolated from each SILAC condition (4/6 and 0/0) and then equal amount of DRM was combined (post fractionation) and then analyzed by mass spectrometry (MS).

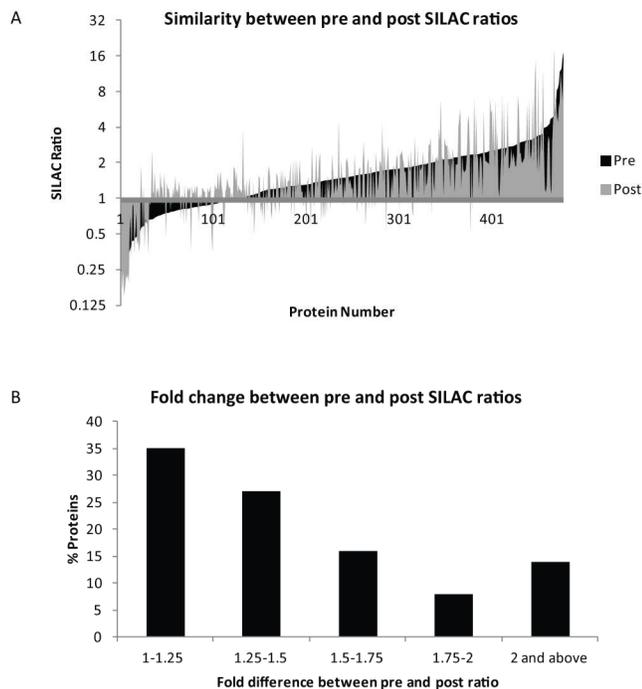


Figure 2. Similarity of SILAC ratios between mixing methods. (A) Pre fractionation mix SILAC ratios were sorted in ascending order and plotted (black). The corresponding post fractionation mix SILAC ratio was overlaid in grey. (B) Bar graph representing the fold difference in SILAC ratios from pre fractionation and post fractionation mix methods.

lated the DRM for LC-MS/MS analysis (referred to as pre fractionation mix) (Figure 1). To compare combining SILAC samples downstream of fractionation, we separately isolated DRM from SILAC labeled WT MEFs and Cav1^{-/-} MEFs then combined equal amounts of DRM prior to LC-MS/MS (referred to as post fractionation mix) (Figure 1).

SILAC ratios were determined by calculating an overall mean for each protein using SILAC ratios for peptides from 3 independent biological replicate experiments (Supplementary Table 1). We examined the similarity of the SILAC ratios produced from the pre fractionation and post fractionation mixing methods for all overlapping protein (Figure 2A). We found that 62% of proteins had ratios within 1.5 fold and 86% displayed less than 2 fold variation (Figure 2B). These data indicate that normalization of protein amount pre fractionation or post fractionation generates similar SILAC ratios for the majority of proteins. However, a small subset of proteins (14%) showed a greater than 2 fold difference in SILAC ratios between mixing methods.

3.2 Analysis of proteins regulated by caveolin-1

To discover proteins affected by the absence of caveolin-1, we analyzed proteins that were significantly different between WT MEFs and Cav1^{-/-} MEFs. Proteins that had a SILAC ratio two or more standard deviations from the mean and a p-value less than 0.05 were considered significantly different (Figure 3A). The pre fractionation method identi-

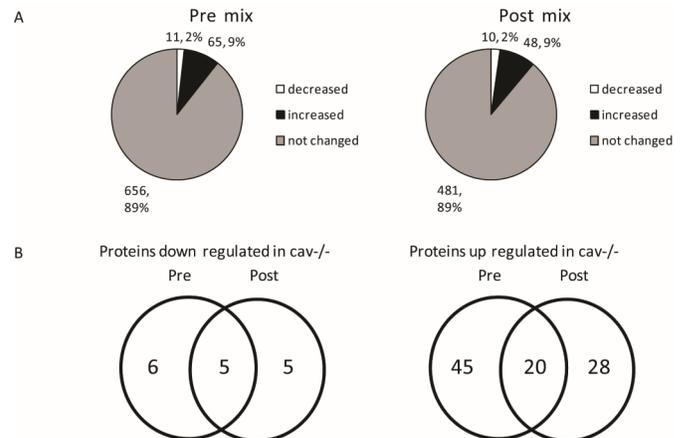


Figure 3. Proteins regulated by caveolin-1 in pre fractionation and post fractionation methods. (A) Pie chart indicating proteins not changed, increased or decreased (mean \pm 2SD) in the pre fractionation and post fractionation mix methods. (B) Overlapping up regulated and down regulated proteins (mean \pm 2SD) from pre fractionation and post fractionation methods is displayed in Venn diagrams.

fied 11 (2%) down regulated proteins and 65 (9%) proteins upregulated from a total of 732 proteins (Figure 3A, left panel). The post fractionation mix method produced very similar percentages with 10 (2%) proteins that were down regulated and 48 (9%) upregulated proteins from the 539 proteins quantified (Figure 3A, right panel). Despite the comparable percentages, only 5 down regulated proteins were common to both mixing methods from the 11 and 10 quantified proteins respectively, and 20 up regulated proteins common from 65 and 45 identified (Figure 3B). Analysis of the overlapping proteins significantly altered by caveolin-1 (Table 1) revealed extensive variation between pre and post fractionation SILAC ratios. This was unexpected since we had shown considerable similarity between SILAC ratios when assessing all quantified proteins. It is likely that the small percentage of proteins that have a greater than 2 fold variation between SILAC ratios are biologically responsive to loss of caveolin-1.

Since the pre fractionation mixing is normalized by total protein amount, one potential explanation of the observed difference is that Cav1^{-/-} MEFs contain a different proportion of DRM compared to WT MEF. To address this possibility, we measured DRM protein amounts from the two cell types and expressed it as a percentage of the starting total cellular protein amount. This experiment determined that the DRM constitutes 0.51 ± 0.06 SEM % of total protein in WT MEFs and 0.84 ± 0.09 SEM % in Cav1^{-/-} MEFs (Figure 4A). This equates to a 1.6 fold increase in protein recruitment to DRM. While biologically interesting, this result has serious implications on the SILAC pre fractionation mixing method since the Cav1^{-/-} MEFs will contribute more protein to the mix than WT MEFs. It is probable the SILAC ratios will be skewed making the data difficult to interpret. To understand how this result could potentially influence our data, we mathematically modeled the effect on the pre and post fractionation mix methods. Since equal levels of DRM protein was combined in the post fractionation method, the

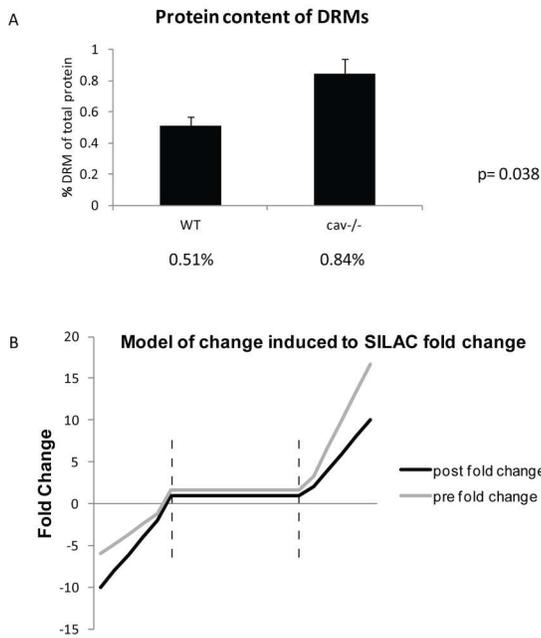


Figure 4. Loss of caveolin-1 increases protein recruitment to DRM. (A) Protein content of DRM and whole cell lysate was determined using Bradford assay. The graph shows the percentage of DRM protein represents from total cellular protein (WT MEFs = 0.51 ± 0.06 SEM % and Cav1^{-/-} MEFs = 0.84 ± 0.09 SEM % of total protein). Using student T-Test, there is a significant difference between WT MEFs and Cav1^{-/-} MEFs, p = 0.038. (B) Model of the affect that increased DRM protein in Cav1^{-/-} MEFs is predicted to have on pre fractionation mixing and post fractionation mixing. The pre fractionation mix method fold changes were adjusted to reflect the predicted 1.6 fold increase contributed by Cav1^{-/-} DRM protein.

Table 1. Overlapping proteins altered by caveolin-1

ID	Protein Name	Pre ratio	Post ratio
Q9JM51	Prostaglandin E synthase	0.27	0.15
P01831	Thy-1 membrane glycoprotein	0.30	0.18
Q9JLI3	Membrane metallo-endopeptidase-like 1	0.32	0.21
P21995	Embigin	0.27	0.25
P97449	Aminopeptidase N	0.24	0.26
P25911	Tyrosine-protein kinase Lyn	8.52	4.39
Q9CYL5	Golgi-associated plant pathogenesis-related protein 1	3.47	4.54
P62071	Ras-related protein R-Ras2	4.93	4.57
Q3UMR5	Coiled-coil domain-containing protein 109A	5.35	4.66
Q9D379	Epoxide hydrolase 1	8.36	4.92
P21956	Lactadherin	9.26	5.86
Q9DCJ5	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8	3.09	6.03
Q9CU24	Protein THEMIS3	11.76	6.48
Q8VHL0	Urea transporter 1	15.27	6.62
Q9WV54	Acid ceramidase	3.12	6.64
Q6X893	Choline transporter-like protein 1	6.59	6.85
P19783	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	3.14	6.91
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	3.25	8.49
Q6ZQM8	UDP-glucuronosyltransferase 1-7C	12.56	8.80
Q9CPQ1	Cytochrome c oxidase subunit 6C	3.42	9.77
Q80Z24	Neuronal growth regulator 1	12.17	11.58
P56391	Cytochrome c oxidase subunit 6B1	3.73	11.67
Q61468	Mesothelin	17.13	13.69
P48771	Cytochrome c oxidase subunit 7A2, mitochondrial	3.25	14.56
Q7SIG6	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2	4.95	17.29

SILAC fold change was left unchanged. This was compared to the pre fractionation mix method, which the SILAC fold change was adjusted by 1.6 fold to account for the predicted increased contribution by Cav1^{-/-} MEFs (Figure 4B). The model calculated a minor increase in the SILAC fold change for pre fractionation mixing compared with post fractionation mixing for proteins that were not altered by caveolin-1 (fold change between the broken lines). However, in stark contrast, proteins that were altered by caveolin-1 (fold change outside of the broken lines) were predicted to show variation between mixing methods that radically increased as the fold change increased. More specifically, our model predicted that protein down regulated in Cav1^{-/-} MEFs (a negative fold change) in the post fractionation mix exhibited an increased SILAC fold change in the pre fractionation mix thereby making it less significant. The opposite trend was predicted for up regulated proteins with an increase in the SILAC fold change with pre fractionation mix compared to post fractionation, making it appear more significant. Consistent with this model, we found the majority of overlap-

ping proteins showed relatively similar SILAC ratios between mix methods, but large differences in proteins altered by caveolin-1 expression. This would have a major impact on whether a protein is considered significantly altered or not. This may also explain the lower percentage of overlap of altered protein we observed between mix methods.

We next examined our list of overlapping altered proteins to see if they behaved as the model predicted. Table 2 shows that 11 of the 25 of the proteins did behave as predicted, a small subset of proteins (5) showed no change in SILAC ratios between mix methods, while 9 proteins showed opposite trend to the model. Given that DRM is derived from membrane from all cellular compartments including plasma membrane, Golgi, and ER, we also assigned the subcellular location using Uniprot to our list of altered proteins (Table 2). Interestingly, the altered proteins that behaved as the model predicted were found to be predominately membrane/cell membrane localized. However, those that behaved in the opposite manner to the model were mainly mitochondria and Golgi localized. These data suggest that there is

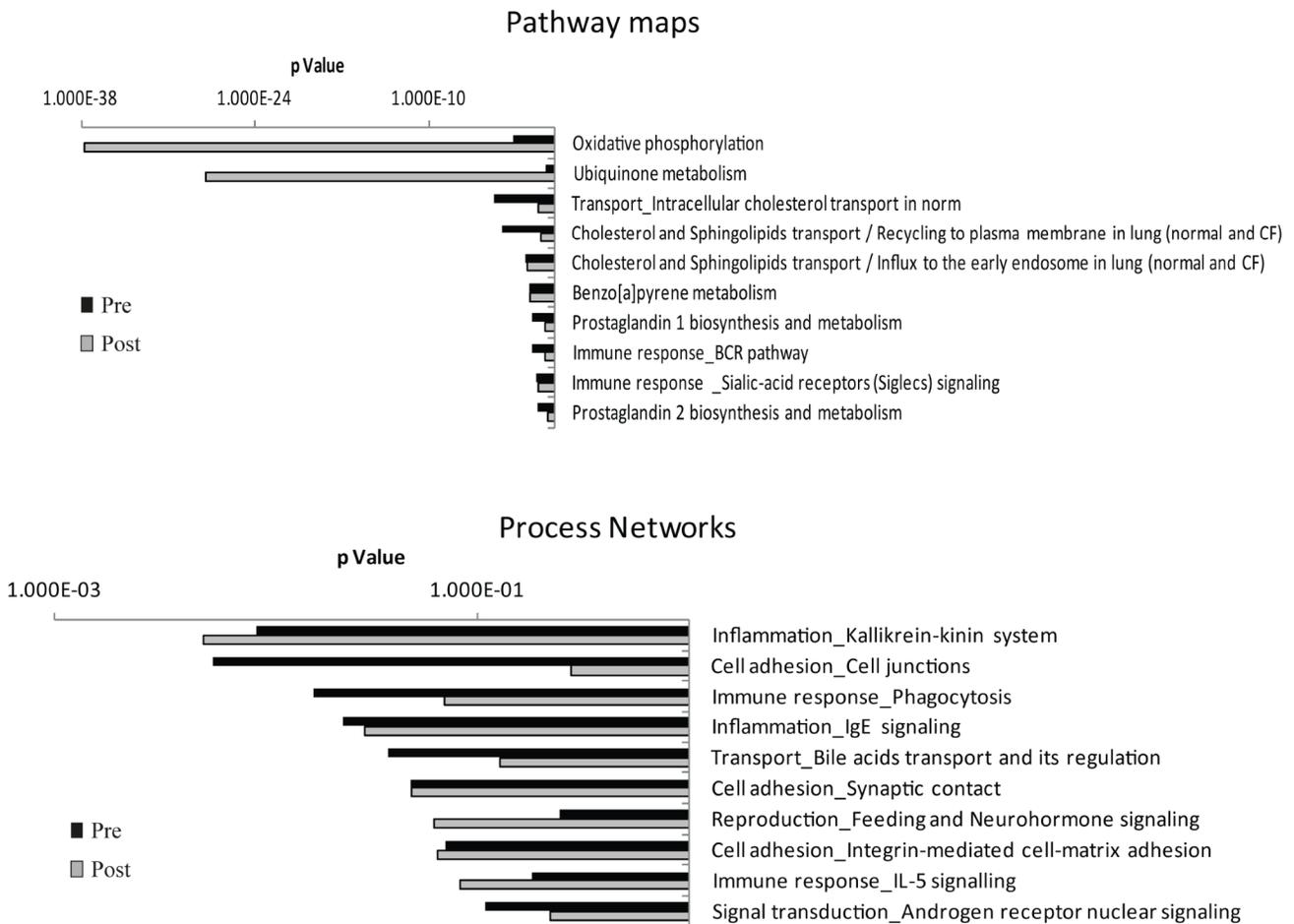


Figure 5. Mixing method impacts on the functional pathways enriched in the significantly altered protein list. Bar chart of GeneGo analysis displaying the most significantly enriched pathways and biological processes for pre fractionation mixing and post fractionation mixing.

Table 2. Proteins altered by caveolin-1, compliance with model and the subcellular localisation.

ID	Protein Name	Behave as modeled	Subcellular Location
P25911	Tyrosine-protein kinase Lyn	Yes	Cell membrane. Nucleus. Cytoplasm. Cytoplasm > perinuclear region. Golgi apparatus.
P01831	Thy-1 membrane glycoprotein	Yes	Cell membrane; Lipid-anchor > GPI-anchor.
Q61468	Mesothelin	Yes	Cell membrane; Lipid-anchor > GPI-anchor. Golgi apparatus. Secreted. Cell membrane; Multi-pass membrane protein. Basolateral cell membrane.
Q8VHL0	Urea transporter 1	Yes	
Q9JM51	Prostaglandin E synthase	Yes	Membrane; Multi-pass membrane protein.
P21956	Lactadherin	Yes	Membrane; Peripheral membrane protein. Secreted.
Q9JLI3	Membrane metallo-endopeptidase-like 1	Yes	Membrane; Single-pass type II membrane protein. Secreted.
Q9D379	Epoxide hydrolase 1	Yes	Microsome membrane; Single-pass type II membrane protein. Endoplasmic reticulum membrane; Single-pass type II membrane protein.
Q6ZQM8	UDP-glucuronosyltransferase 1-7C	Yes	Microsome. Endoplasmic reticulum membrane; Single-pass membrane protein.
Q3UMR5	Coiled-coil domain-containing protein 109A	Yes	Mitochondrion inner membrane; Multi-pass membrane protein.
Q9CU24	Protein THEMIS3	Yes	unknown
Q80Z24	Neuronal growth regulator 1	No change	Cell membrane; Lipid-anchor > GPI-anchor.
P62071	Ras-related protein R-Ras2	No change	Cell membrane; Lipid-anchor; Cytoplasmic side. Cell membrane; Multi-pass membrane protein. Mitochondrion outer membrane; Multi-pass membrane protein.
Q6X893	Choline transporter-like protein 1	No change	
P21995	Embigin	No change	Membrane; Single-pass type I membrane protein.
P97449	Aminopeptidase N	No change	Membrane; Single-pass type II membrane protein.
Q7SIG6	Arf-GAP with SH3 domain-containing protein 2	No	Golgi apparatus > Golgi stack membrane; Peripheral membrane protein. Cell membrane; Peripheral membrane protein. Cytoplasm
Q9CYL5	Golgi-associated plant pathogenesis-related protein 1	No	Golgi apparatus membrane; Lipid-anchor.
Q9WV54	Acid ceramidase	No	Lysosome.
P19783	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	No	Mitochondrion inner membrane.
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	No	Mitochondrion inner membrane.
Q9CPQ1	Cytochrome c oxidase subunit 6C	No	Mitochondrion inner membrane.
P48771	Cytochrome c oxidase subunit 7A2, mitochondrial	No	Mitochondrion inner membrane.
P56391	Cytochrome c oxidase subunit 6B1	No	Mitochondrion intermembrane space.
Q9DCJ5	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8	No	Mitochondrion. Mitochondrion intermembrane space.

probably not a global up regulation of all DRM proteins in Cav1-/- MEFs. Rather, caveolin-1 has distinct effects on DRMs from different subcellular compartments. Although caveolin-1 is critical for plasma membrane caveolae, it has also been identified at multiple other locations with no morphological caveolae, including the Golgi [17], endosomes [18] and the mitochondria [19]. However the role of non-caveolar caveolin remains unclear.

3.3 Functional analysis of caveolin-1 regulated proteins

To discover how the pre and post fractionation mix methods impact upon functional analysis of proteomics data, we performed pathways enrichment analysis using GeneGo software on the proteins significantly altered by caveolin-1 from both mixing methods (Figure 5). Post fractionation mixing data suggested that caveolin-1 regulated DRM proteins involved in oxidative phosphorylation and ubiquitone

metabolism (Figure 5, top panel) however, these pathways were far less significant with the pre fractionation mix method. In addition, pre and post fractionation mix methods revealed caveolin-1 regulated inflammation but cell adhesion was dramatically over represented in the pre fractionation mix method (Figure 5, bottom panel). Therefore, these results clearly demonstrate that different conclusions are derived when mixing pre or post fractionation.

Specific quantitative differences in the mitochondria DRM proteins were observed between pre and post fractionation mixing set up (Table 2). The post fractionation mix method highlighted a role for caveolin-1 in regulation of mitochondrial proteins and oxidative phosphorylation. Identification of proteins annotated as mitochondrial proteins in the DRM have been reported in several proteomics studies, and a number of different methods have been used to determine if these proteins were contaminants during biochemical fractionation [21-23]. Foster *et al.* differentiated true DRM proteins from contaminants using sensitivity to cholesterol disruption. In this scenario, only *bona fide* DRM proteins are responsive to changes in cholesterol [20]. Follow up studies with this method found that the major mitochondrial proteins reported in proteomics studies such as ATP synthase subunits and voltage-dependent anion selective channels (VDACs) were DRM contaminants [21]. In stark contrast, a recent study demonstrated the opposite result with certain mitochondrial proteins sensitive to cholesterol disruption [22]. Elegant studies using affinity purification of cell surface lipid rafts, a method optimized to exclude mitochondrial contamination, reported the presence of mitochondrial proteins ATP synthase and cytochrome c oxidase [23]. Therefore it remains highly controversial if mitochondrial proteins are true DRM proteins, a topic reviewed extensively elsewhere [23, 24]. In our study, the observed consistent quantitative differences in mitochondrial proteins points to a true caveolin-1 effect. Further support of this comes from recent studies showing that caveolin-1 participates in mitochondrial cholesterol regulation [25, 26]. In the absence of caveolin-1, cholesterol was found to congregate in the membrane of mitochondria, affecting the function of mitochondria [26]. Importantly, this finding would not have been revealed without comparison of pre and post fractionation methods, and examination of relative DRM amounts in the cell types.

4. Concluding Remarks

Finally, our results indicate that careful consideration is needed when deciding on what step is best to combine SILAC samples. This will vary with each experimental model and we therefore recommend a preliminary examination of how experimental conditions affect the organelle/s of interest before proceeding with SILAC subcellular proteomics in order to make sure correct interpretation can be made.

5. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/108/0>.

Supplementary Material includes Supplementary Table 1. SILAC ratios and p values for all proteins identified.

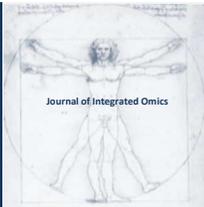
Acknowledgements

We thank Eunju Choi for technical assistance and useful discussions. MMH is supported by a National Health and Medical Research Council of Australia (NHMRC) Career Development Fellowship (App ID 569512). Access to GeneGo was provided by the Queensland Facility for Advanced Bioinformatics (QFAB) through LIEF Grant LE098933.

References

1. K. Inder, A. Harding, S. J. Plowman, M. R. Philips, R. G. Parton, and J. F. Hancock, *Mol Biol Cell* 19 (2008) 4776-4784.
2. K. L. Inder, C. Lau, D. Loo, N. Chaudhary, A. Goodall, S. Martin, A. Jones, D. van der Hoeven, R. G. Parton, M. M. Hill, and J. F. Hancock, *J Biol Chem* 284 (2009) 28410-28419.
3. S. Brunet, P. Thibault, E. Gagnon, P. Kearney, J. J. Bergeron, and M. Desjardins, *Trends in cell biology* 13 (2003) 629-638.
4. J. R. Yates, 3rd, A. Gilchrist, K. E. Howell, and J. J. Bergeron, *Nature reviews. Molecular cell biology* 6 (2005) 702-714.
5. J. S. Andersen, Y. W. Lam, A. K. Leung, S. E. Ong, C. E. Lyon, A. I. Lamond, and M. Mann, *Nature* 433 (2005) 77-83.
6. V. K. Mootha, J. Bunkenborg, J. V. Olsen, M. Hjerrild, J. R. Wisniewski, E. Stahl, M. S. Bolouri, H. N. Ray, S. Sihag, M. Kamal, N. Patterson, E. S. Lander, and M. Mann, *Cell* 115 (2003) 629-640.
7. S. W. Taylor, E. Fahy, B. Zhang, G. M. Glenn, D. E. Warnock, S. Wiley, A. N. Murphy, S. P. Gaucher, R. A. Capaldi, B. W. Gibson, and S. S. Ghosh, *Nature biotechnology* 21 (2003) 281-286.
8. C. E. Au, A. W. Bell, A. Gilchrist, J. Hiding, T. Nilsson, and J. J. Bergeron, *Curr Opin Cell Biol* 19 (2007) 376-385.
9. S. D. Patterson and R. H. Aebersold, *Nat Genet* 33 Suppl (2003) 311-323.
10. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann, *Mol Cell Proteomics* 1 (2002) 376-386.
11. E. Emmott, M. A. Rodgers, A. Macdonald, S. McCrory, P. Ajuh, and J. A. Hiscox, *Mol Cell Proteomics* 9 (2010) 1920-1936.
12. S. Dhungana, B. A. Merrick, K. B. Tomer, and M. B. Fessler, *Molecular & cellular proteomics : MCP* 8 (2009) 201-213.
13. K. L. Inder, Y. Z. Zheng, M. J. Davis, H. Moon, D. Loo, H. Nguyen, J. A. Clements, R. G. Parton, L. J. Foster, and M. M. Hill, *Molecular & cellular proteomics : MCP* 11 (2012) M111012245.
14. R. G. Parton and K. Simons, *Nat Rev Mol Cell Biol* 8 (2007) 185-194.
15. M. M. Hill, M. Bastiani, R. Luetterforst, M. Kirkham, A. Kirkham, S. J. Nixon, P. Walser, D. Abankwa, V. M. Oorschot, S. Martin, J. F. Hancock, and R. G. Parton, *Cell* 132 (2008) 113-124.
16. M. Kirkham, A. Fujita, R. Chadda, S. J. Nixon, T. V. Kurzchalia, D. K. Sharma, R. E. Pagano, J. F. Hancock, S. Mayor, and R. G. Parton, *J Cell Biol* 168 (2005) 465-476.

17. T. V. Kurzchalia and R. G. Parton, *Current opinion in cell biology* 11 (1999) 424-431.
18. A. Pol, M. Calvo, A. Lu, and C. Enrich, *Hepatology* 29 (1999) 1848-1857.
19. W. P. Li, P. Liu, B. K. Pilcher, and R. G. Anderson, *Journal of cell science* 114 (2001) 1397-1408.
20. M. Bosch, M. Mari, S. P. Gross, J. C. Fernandez-Checa, and A. Pol, *Traffic* 12 (2011) 1483-1489.
21. M. Bosch, M. Mari, A. Herms, A. Fernandez, A. Fajardo, A. Kassan, A. Giralt, A. Colell, D. Balgoma, E. Barbero, E. Gonzalez-Moreno, N. Matias, F. Tebar, J. Balsinde, M. Camps, C. Enrich, S. P. Gross, C. Garcia-Ruiz, E. Perez-Navarro, J. C. Fernandez-Checa, and A. Pol, *Current biology : CB* 21 (2011) 681-686.
22. K. B. Kim, J. W. Lee, C. S. Lee, B. W. Kim, H. J. Choo, S. Y. Jung, S. G. Chi, Y. S. Yoon, G. Yoon, and Y. G. Ko, *Proteomics* 6 (2006) 2444-2453.
23. B. W. Kim, C. S. Lee, J. S. Yi, J. H. Lee, J. W. Lee, H. J. Choo, S. Y. Jung, M. S. Kim, S. W. Lee, M. S. Lee, G. Yoon, and Y. G. Ko, *Expert review of proteomics* 7 (2010) 849-866.
24. L. J. Foster and Q. W. Chan, *Sub-cellular biochemistry* 43 (2007) 35-47.
25. M. Bosch, M. Mari, S. P. Gross, J. C. Fernandez-Checa, and A. Pol, *Traffic* 12 (2011) 1483-1489.
26. M. Bosch, M. Mari, A. Herms, A. Fernandez, A. Fajardo, A. Kassan, A. Giralt, A. Colell, D. Balgoma, E. Barbero, E. Gonzalez-Moreno, N. Matias, F. Tebar, J. Balsinde, M. Camps, C. Enrich, S. P. Gross, C. Garcia-Ruiz, E. Perez-Navarro, J. C. Fernandez-Checa, and A. Pol, *Current biology : CB* 21 (2011) 681-686.



JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.112

Proteomic analysis of bronchoalveolar lavage fluid in an equine model of asthma during a natural antigen exposure trial

Marybeth Miskovic Feutz¹, C. Paige Riley², Xiang Zhang³, Jiri Adamec⁴, Craig Thompson⁵, Laurent L. Couetil^{*6}

1Purdue University College of Veterinary Medicine, Purdue University, 625 Harrison St., West Lafayette, Indiana, 47907, United States of America; 2Bindley Bioscience Center, Purdue University, 1203 W. State St., West Lafayette, Indiana, 47907, United States of America; 3University of Louisville, 2320 S. Brook St., Louisville, Kentucky, 40292, United States of America; 4Bindley Bioscience Center, Purdue University, 1203 W. State St., West Lafayette, Indiana, 47670, United States of America; 5Purdue University College of Veterinary Medicine, Purdue University, 625 Harrison St., West Lafayette, Indiana, 47907, United States of America; 6Purdue University College of Veterinary Medicine, Purdue University, 625 Harrison St., West Lafayette, Indiana, 47907, United States of America.

Received: 21 September 2012 Accepted: 14 November 2012 Available Online: 08 December 2012

ABSTRACT

Background

Heaves is a complex, asthma-like respiratory disease that affects many older horses. While environmental and genetic components to the disease have been proposed, the specific pathophysiology of heaves is still poorly understood. Using proteomic techniques, we compared the protein profile of bronchoalveolar lavage fluid (BALF) in the lungs of healthy horses and horses affected with heaves.

Methods

Clinical signs of the disease were induced in heaves-affected horses using an experimental hay exposure model. Samples of BALF were collected from all horses before and after the hay exposure trial. Mass spectrometry (LC-MS) was used to evaluate the differences in the global BALF peptide profile between the control and heaves-affected horses. Tandem mass spectrometry (LC-MS/MS) was used to identify differentially expressed proteins in the two groups of horses. The identification of two proteins was validated with Western blot assays.

Results

One hundred peptides were differentially expressed between healthy controls and heaves-affected horses; 76 peptides were over-expressed in controls and 24 were over-expressed in heaves-affected horses. The identifications of transferrin and secretoglobin were confirmed with Western blot.

Conclusions

This study demonstrates that proteomics can be used to compare the protein profiles of BALF from healthy and diseased horses. These techniques may prove helpful in determining the pathophysiology of complex diseases.

Keywords: Bronchoalveolar lavage fluid; Transferrin; Secretoglobin; Heaves.

Abbreviations:

BAL, bronchoalveolar lavage; **BALF**, bronchoalveolar lavage fluid; **BSA**, bovine serum albumin; **C_{dyn}**, dynamic lung compliance; **ΔP_L**, transpulmonary pressure; **ΔP_{Lmax}**, maximum change in transpulmonary pressure; **LC-MS**, liquid chromatography-mass spectrometry; **LC-MS/MS**, liquid chromatography-mass spectrometry/mass spectrometry (tandem mass spectrometry); **NHS**, normal horse serum; **PDP**, Purdue Discovery Pipeline; **PFT**, pulmonary function testing; **R_L**, total lung resistance; **TBS**, Tris-buffered saline.

*Corresponding author: Laurent L. Couetil, Purdue University College of Veterinary Medicine, Purdue University, 625 Harrison St., West Lafayette, Indiana, 47907, United States of America; Phone number: 1-765-494-8548; Fax Number: 1-765-496-2641; E-mail address: couetil@purdue.edu

1. Introduction

The prevalence of heaves has been reported to vary between 3-20% [1-3]. Heaves is an asthma-like disease of horses that typically manifests with clinical signs such as increased respiratory rate and effort, productive cough, and exercise intolerance. This reaction appears to be an allergic response to the molds and dusts present in the hay, straw, and barn environment, and results from prolonged exposure to environments that contain high amount of aeroallergens [4]. A genetic predisposition for heaves is supported by several reports. In one report, the prevalence of heaves in two populations of Swiss warmblood horses was higher in offsprings from two heaves-affected stallions compared to control populations from non-heaves-affected stallions.⁵ In this population of horses, an association was found between clinical signs of heaves and microsatellite markers near the gene for the IL-4 receptor in the descendants of one heaves-affected Swiss warmblood stallion, but not the second stallion, suggesting that this is a multifactorial, perhaps polygenic, disease [5].

Since heaves is triggered by exposure to inhaled allergens, the analysis of bronchoalveolar lavage fluid (BALF) would seem to be the most likely fluid to yield potential biomarkers for disease. However, over 500 proteins have been identified in equine BALF making detection of differential protein expression using traditional methods (e.g. ELISA, Western blot) a daunting task [6]. In order to evaluate a larger number of proteins at once, we used proteomic techniques during the discovery phase of this experiment. Proteomics uses large-scale protein expression technologies for the identification and quantification of proteins that might be altered in response to various insults or diseases [7]. As a result, proteomics can be very valuable to investigate the pathophysiology of complex diseases.

Although the proteome of equine BALF has previously been reported in healthy horses [6], BALF proteome changes in heaves have not been reported. Our hypothesis was that the BALF of heaves-affected horses would show differential protein expression, compared to the BALF of healthy horses. We chose to use LC-MS and LC-MS/MS protocols as this approach allows for global evaluation of average relative peptide concentrations between groups of horses in the first step of analysis (LC-MS) and identification of highly expressed peptides in the second step of analysis (LC-MS/MS) [8,9]. Finally, we used Western blot techniques to confirm the identity and expression pattern of two proteins identified with proteomic techniques.

2. Material and Methods

Animals. Horses with heaves ($n = 8$) were selected from a herd of horses owned by Purdue University College of Veterinary Medicine (PUCVM). These horses had been previously diagnosed with heaves based on maximum

change in transpulmonary pressure (ΔP_{Lmax}) ≥ 15 cmH₂O, reversible airway obstruction, and $\geq 25\%$ neutrophils in BALF cytology during disease exacerbation [10]. Control horses ($n = 8$) were recruited from the PUCVM teaching herd. The control horses were judged to be normal based on physical examination and no history of clinical signs attributable to chronic respiratory disease when fed hay and housed indoors. The heaves-affected horses were 4 females and 4 castrated males; 5 Quarter Horse-type breeds, 1 walking horse, 1 Arabian cross, and 1 Icelandic pony. The control horses were 7 females and 1 castrated male; 3 standardbreds, 2 Quarter Horse-type breeds, 1 thoroughbred, 1 walking horse, and 1 Arabian. The Purdue University Animal Care and Use Committee approved all procedures.

All horses were maintained on pasture for at least two months before the beginning of this study. On Day 1, all horses were transported from the pasture to the laboratory and allowed at least 30 minutes of acclimatization. The evaluation included a complete physical examination, standard pulmonary function testing (PFT), and BAL. After the horses recovered from sedation, they were returned to the pasture for one week. This testing was repeated three times over two weeks to assess the stability of the measurements over time. After the third test, each pair of horses (one heaves-affected and one control) was stalled in a barn in adjacent stalls for a hay exposure trial.

Experimental Exposure. All horses were fed moldy hay (0.2 kg/45 kg body weight) mixed with good quality hay (0.5 kg/45 kg body weight) and pelleted feed^a (0.3 kg/45 kg body weight), and were bedded on straw. A physical examination was performed and a clinical score based on severity of clinical signs was calculated daily on each horse. The clinical score is based on the respiratory rate and effort, presence of abnormal lung sounds, and degree of nasal discharge [11]. Horses were exposed to hay for up to 21 days, or until the clinical score of the heaves-affected horse reached 10 (out of 21 possible). When the heaves-affected horse had a clinical score of ≥ 10 , PFT was performed; when the heaves-affected horse had a $\Delta P_{Lmax} \geq 15$ cmH₂O, the tests performed at baseline were repeated on the heaves-affected horse and its control.

Pulmonary Function Evaluation. Testing was performed according to standard laboratory procedure, as previously described [12]. Horses were restrained in stocks without sedation to allow breath-by-breath measurement of esophageal and mask pressures as well as airflow. Data analysis of ten representative breaths yielded total pulmonary resistance, dynamic compliance, and ΔP_{Lmax} . After completion of the PFT, horses were sedated with detomidine^b (0.03 mg/kg) and butorphanol^c (0.02 mg/kg) and BAL was performed as previously reported, using a flexible 2-meter videoendoscope [12].^d Two different segments of the same lung were infused with 250 mL sterile saline per site. The aspirated BALF was pooled, immediately placed on ice, and processed within 20 minutes of collection. An aliquot was prepared for cytological examination by

cyto centrifugation and stained with modified Wright's stain. Total nucleated cell counts were determined by use of a hemacytometer. Differential cell counts were determined by examination of 200 leukocytes per slide. The clinical pathologist evaluating the BALF cytology was blinded to each horse's group identity.

Sample Preparation for Proteomics Analysis. The BALF was prepared for proteomic analysis according to standard laboratory procedures (details provided in Additional File 1). Briefly, the BALF was filtered through sterile gauze and centrifuged. Protein concentration of the supernatant was determined with a BCA Assay^e and a volume equivalent to 100 µg of protein from each BALF sample was incubated with cold acetone at -20 °C to precipitate proteins. The precipitate was lyophilized and digested with trypsin. Samples were processed in batches, and the peptide solutions were stored at -20 °C until mass spectrometric analysis.

LC-MS Analysis. In this study, a two-step proteomic approach was used. In the first step, all available post-exposure samples (heaves-affected n = 5, control n = 6) were analyzed by a simple LC-MS technique for quantitative purposes. This data was used to generate the average relative peptide concentrations within each group of horses (heaves-affected and controls), and to determine the fold change between the two groups. In second step we performed MS/MS analysis on BALF samples from one heaves-affected and one control horse.

Peptides from each BALF sample at the exposure testing period (control n = 6; heaves-affected n = 5; 1 µg of peptides from each sample) were subjected to LC-MS analysis on a nanoLC-Chip system,^f according to standard laboratory protocol (details in Additional File 1). Buffer A (5% ACN/0.01% TFA) was used to concentrate the peptides on the on-chip enrichment column and to separate the peptides on the reversed phase analytical column. The column was eluted with a 55 minute gradient from 5%-35% buffer B (100% ACN/0.01% TFA), followed by a 10 minute gradient from 35%-100% buffer B. The column was re-equilibrated between each sample with isocratic flow of 5% buffer B. The system was controlled by ChemStation software.^g

Data generated with LC-MS were analyzed with the Purdue Discovery Pipeline^h (PDP). Average relative concentrations of each peak in the groups (heaves-affected and controls) and the relative fold change for each peak between groups were determined [13]. The raw data from the LC-MS were pre-processed before analysis to eliminate artifacts such as noise, peak broadening, instrument distortion, etc [13]. The mean intensity of each peak was calculated for each group (heaves-affected n = 5 and control n = 6) and the fold change of each peptide between groups was calculated. When a peptide was absent in a group of horses, a placeholder value of "10" was used to facilitate statistical analysis. A peak had to be found in at least 7 of 11 BALF samples or in at least 1 of the BALF samples from one group (heaves-affected or control) to be considered in the analysis.

LC-MS/MS Analysis. In a separate experiment, BALF samples from one heaves-affected and one control horse from the exposure testing period were subjected to LC-MS/MS (tandem MS) analysis on the same nanoLC-chip system^f according to standard laboratory procedure (details in Additional File 1). Assuming each peptide is present in every heaves-affected and control horse, analyzing a single sample should be sufficient for identification purposes. But because not every peptide was present in both groups of horses, we chose a single representative sample from each group for LC-MS/MS analysis based on the LC-MS peak list of those samples to ensure that 1) most of the significantly different peptide peaks were detected in that sample, and 2) the peptide peak of the regulated peptides were intense peaks. Automated MS/MS spectra were acquired during each run in the data-dependent acquisition mode, with the selection of the three most abundant precursor ions in each peak. Analysis of the data generated during tandem MS runs yielded identifications of the parent proteins of the peptides detected (see Peak Identification, below).

Peak Identification. Files acquired during the tandem MS runs were uploaded to Spectrum Mill protein identification software.^g Two databases were downloaded from NCBIⁱ in 2009, the non-redundant equine protein database (downloaded 2/24/09; 29,374 proteins) and the non-redundant mammalian database (downloaded 1/26/09; 7,745,744 proteins). Identification of the peptides was performed using Spectrum Mill to separately search the two NCBI protein databases. The search parameters were: no more than two tryptic miscleavages allowed, cysteine searched as ethanol cysteine, variable oxidized methionine, 2.5 Da peptide tolerance, and 0.7 Da mass tolerance. Only peptides with a score of 5 or higher were considered true positives. The identification data from Spectrum Mill analysis was matched with the concentration data from PDP analysis by merging the output files from the two data analyses for each horse into one file. A list of peptide identifications, with associated relative concentrations, was made for as many peaks as possible, based on m/z, retention time, and peptide charge. Please see Additional File 1 for more details.

Western blot. Western blots were performed according to the manufacturer's recommendations^j on all baseline and exposure BALF samples (controls n = 8 for baselines 1, 2, 3 and exposure; heaves-affected n = 5 for baselines 1, 2, exposure and n = 4 for baseline 3) with 2 µg of protein from each sample. Ten BALF samples were concentrated to achieve a protein concentration of approximately 0.2-0.3 µg/µL.^k ImageJ software^l was used to quantify the density of the bands from the protein of interest on each Western blot. (Details in Additional File 1)

Secretoglobin. In order to validate the identification of secretoglobin from the proteomics data, Western blot was performed on all baseline and exposure BALF samples. Two µg of protein from each sample was prepared as described above. The positive control protein was equine recombinant secretoglobin.^l The primary antibody was monoclonal rabbit

anti-horse secretoglobulin antiserum [14]^l, at a dilution of 1:1000. The secondary antibody was horseradish peroxidase-conjugated goat anti-rabbit IgG,^m at a dilution of 1:2000.

Transferrin. In order to validate the identification of transferrin from the proteomics data, Western blot was performed on all baseline and exposure BALF samples. Two μ g of protein from each sample was prepared as described above. The positive control protein was human apo-transferrin.ⁿ The primary antibody was polyclonal sheep anti-horse transferrin,^o at a dilution of 1:10,000. According to the manufacturer, this antibody is specific for equine transferrin. The secondary antibody was horseradish peroxidase-conjugated donkey anti-sheep IgG (H+L)^m at a dilution of 1:100,000.

Statistical Analysis. After 21 days in the exposure environment, three heaves-affected horses had not reached a clinical score of ≥ 10 with a ΔP_{Lmax} of ≥ 15 cmH₂O. These horses were not included in the rest of this study and comparisons were performed with two groups of horses: controls (baseline n = 8, exposure n = 6) and heaves-affected (baseline and exposure n = 5). Statistical analysis was performed with commercial software.^p All data were tested for normality with the Shapiro-Wilk W test; most of the data were not normally distributed. The Kruskal-Wallis ANOVA was used to assess for a change over the three baseline measurements within each group (control and heaves-affected horses). The variables of interest (clinical score, PFT, BAL) did not show changes over the three baseline measurements therefore, measurements from each horse were averaged, and the averaged values were used for comparison with the exposure measurements.

Because there is not a robust nonparametric test for evaluating changes due to disease, time, and their interaction, factorial ANOVA was used to compare the clinical score, PFT, and BAL variables between the control and heaves-affected horses at the baseline and exposure measurements. When a statistically significant difference was detected with the ANOVA ($p \leq 0.05$), Tukey's post-hoc test was used for pairwise comparisons.

The pairwise two-tailed t-test was used to evaluate the relative concentration difference of a peak (peptide) between sample (control and heaves-affected horses) groups. To con-

trol the false discovery rate (FDR), a peptide with a q-value of ≤ 0.1 and a fold change of ≥ 10 was considered as a peptide with significant concentration difference between two sample groups.

The Western blot baseline measurements were also evaluated for a change over time within each group (control and heaves-affected horses) with the Kruskal-Wallis ANOVA. Neither group showed a significant difference over baseline time points, therefore baseline measurements were averaged for each horse, and the averaged values were used for comparison to the exposure measurements with the factorial ANOVA. The relationship between the density of the bands of secretoglobulin and transferrin and the pulmonary function and BALF cytology variables were evaluated using the Spearman Rank correlation test. A p-value of ≤ 0.05 was considered significant. Data are presented as median [range].

3. Results

Animals. The heaves-affected horses were older (21 [17-26] years) than the controls (15.5 [10-22] years; $p = 0.028$). There was no significant difference in heights or weights between the groups.

Experimental exposure. The heaves-affected horses responded to exposure within 3-7 days and showed a significant increase in clinical score, compared to the control horses (Table 1). After the experimental exposure to hay, the heaves-affected horses showed evidence of airway obstruction characterized by increased ΔP_{Lmax} and R_L and decreased C_{dyn} compared to baseline values and to controls (Table 1).

Bronchoalveolar lavage. After exposure, the percentage of neutrophils significantly increased in both groups compared to baseline but the heaves-affected horses had a significantly higher percentage of neutrophils in BALF than the controls (Table 1).

Protein identification. A total of 2049 chromatographic peaks (peptides) were found in the LC-MS data; of which 370 peptides were identified for a total of 250 unique proteins with high confidence (Additional File 2). Only 43 of these proteins have been identified in the horse previously [6]. Of the 370 identified peptides, 33 were differentially ex-

Table 1. Summary of pulmonary function testing data before (Baseline = average of three baseline measurements) and after (Exposure) experimental exposure trial. Data displayed as median [range]. *significantly different from baseline data; †significantly different from controls at same time point; $p < 0.05$.

	Baseline		Exposure	
	Control, n = 8	Heaves-affected, n = 5	Control, n = 6	Heaves-affected, n = 5
Clinical score	1.2 [0.7-1.7]	1.7 [1.0-2.0]	2.5 [1.0-3.0]	11.0 [7.0-13.0]*†
ΔP_{Lmax}	6.6 [3.1-7.9]	4.3 [3.9-12.1]	6.3 [5.4-8.5]	26.6 [21.8-34.8]*†
C_{dyn}	3.9 [2.0-6.7]	3.6 [1.0-5.7]	3.9 [1.3-5.5]	0.6 [0.4-1.0]*†
R_L	0.53 [0.19-1.09]	0.43 [0.18-1.20]	0.50 [0.41-0.73]	1.83 [1.53-2.08]*†
% neutrophils	3.3 (1.7-7.7)	17.0 (2.7-23.7)	23.0 (15.0-31.0)*	49.0 (30.5-70.0)*†

ΔP_{Lmax} = maximum change in transpulmonary pressure; C_{dyn} = dynamic lung compliance; R_L = total lung resistance

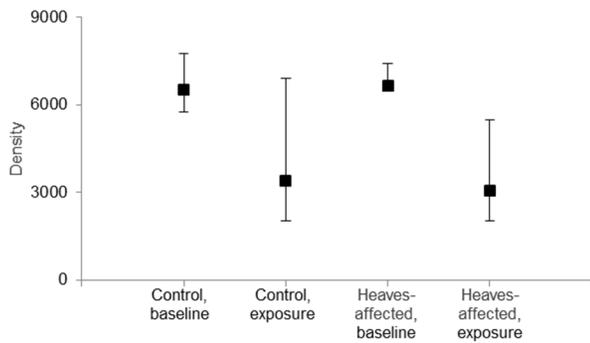


Figure 1. Secretoglobulin expression in bronchoalveolar lavage fluid from control and heaves-affected horses before and after exposure to hay. Area under the curve (density) of transferrin expression on Western blot is displayed as median [Q1, Q3].

pressed with 10 over-expressed in the heaves-affected horses and 23 over-expressed in the control horses (Table 2). Five peptides were identified from secretoglobulin precursor, covering 33% of the protein, including one peptide only present in the precursor protein (peptide KSPLCA, Additional File 3). These peptides did not display differential expression between the heaves-affected and control horses. Fourteen peptides from transferrin precursor protein were identified, covering 24.5% of the protein (Additional File 4); one peptide was absent in the control horses ($p = 0.0094$) and one peptide was absent in heaves-affected horses ($p = 0.012$).

Western blot. Secretoglobulin band density was not significantly different between the heaves-affected and control horses at any time point (Figure 1). There was a trend for secretoglobulin band density during exposure to be lower than at baseline ($p = 0.067$). Transferrin band density was significantly lower in heaves-affected horses at exposure than control horses at baseline (Figure 2; $p = 0.01$). Secretoglobulin band density was negatively correlated with BALF total nucleated ($r = -0.37$, $p = 0.0084$) and neutrophil ($r = -0.40$, $p = 0.0048$) cell counts and positively correlated with the percent of macrophages ($r = 0.41$, $p = 0.0039$; Additional File 5). Transferrin band density was inversely correlated with BALF

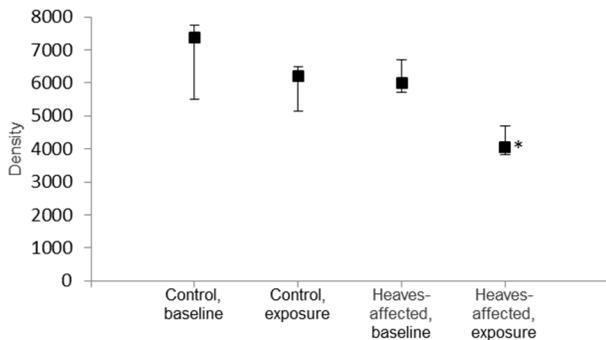


Figure 2. Transferrin expression in bronchoalveolar lavage fluid from control and heaves-affected horses before and after exposure to hay. Area under the curve (density) of transferrin expression on Western blot is displayed as median [Q1, Q3]. *significantly different from controls at baseline, $p = 0.01$.

neutrophil cell count ($r = -0.32$, $p = 0.024$) and positively associated with eosinophilic cell count ($r = 0.32$, $p = 0.023$; Additional File 5).

4. Discussion

The primary finding of this study was the identification of 190 peptides in BALF that were differentially expressed between control and heaves-affected horses after induction of the disease by exposure to moldy hay. This disease model is commonly used to study the pathophysiology of heaves. The three horses that did not respond to the environmental challenge within 21 days in this study had been previously diagnosed with heaves in our laboratory, and had previously responded to a moldy hay challenge. It is possible that these horses were not sensitive to the allergens present in the different batch of moldy hay used in this study. Alternatively, the horses may have shown clinical signs if the exposure trial had continued longer although, most studies report successful exacerbation of heaves after 1-14 days of exposure to moldy hay [16-19].

Results from this study may have been confounded by factors such as age, gender, and breed differences between groups. Horses older than 7 years old are 6.6 times more likely to develop heaves than horses less than 4 years old but this likelihood does not change with increasing age. Horses 7-10 years of age have a similar likelihood of developing heaves as horses 10-15 years of age or horses older than 15 years of age [15]. Therefore, the differences in protein expression in BALF between diseases and control horses are unlikely to be explained by the age difference between the control horses (15.1 [10-22] years) and the heaves-affected horses (21 [17-26] years). Genders were not evenly distributed between groups, however, the risk of heaves is not significantly different between female and castrated horses as compared to intact males [15]. Three of the six breeds represented in the study horses were present in both heaves and control groups, however, two other breeds with an identified risk of heaves (Thoroughbred and Standardbred; [15]) were only represented in the control group. This imbalance may have partially influenced the differential expression of proteins in BALF. Future studies will be needed to tease out the influence of horse signalment (e.g. age, gender, and breed) on differential expression of BALF proteins in horses with heaves.

Lung mechanics measurements at baseline confirmed that none of the horses had detectable airway obstruction with this test, although one heaves-affected horse did have airway neutrophilia. After experimental exposure, as expected, heaves-affected horses showed significant airway obstruction while the control horses did not. Both groups of horses showed evidence of airway inflammation in BALF after the experimental exposure, with a greater increase of neutrophil percentage in the heaves-affected horses. Although the control horses showed no evidence of clinical signs of pulmonary disease and no airway obstruction, these

Table 2. Identified bronchoalveolar lavage fluid proteins that were differentially expressed (q -value ≤ 0.1 and fold change ≥ 10) between heaves-affected and healthy control horses after a hay exposure trial. A positive fold change indicates over-expression in control horses; a negative fold change indicates over-expression in heaves-affected horses. Bold text indicates proteins previously identified in equine BALF.(6)

RefSeq or GenBank #	Fold Change	q-value	Species	Peptide Sequence	Gene Name
AAP80145	10.0	0.042	Equus caballus	(K)EPLFQPQVHVLPVPPSEELALNELVTLTCLVR(G)	immunoglobulin alpha constant heavy chain
CAC86340	-15.6	0.042	Equus caballus	(K)LSVETSR(W)	immunoglobulin gamma 5 heavy chain constant region
CAA46170	10.4	0.042	Rattus norvegicus	(R)GQAPRVTFPGR(G)	integrin alpha chain
EDL02646	-10.0	0.042	Mus musculus	(K)GGGHTTGAGAVSHSAKmWAR(H)	mCG145719
BAC97986	10.1	0.042	Mus musculus	(-)PAPPTPVDRPSSRSVADQR(A)	Arhgap26 Rho GTPase activating protein 26
XP_001917502	10.1	0.042	Equus caballus	(R)NLVYGKANDDNK(I)	catenin (cadherin-associated protein), delta 2
XP_001915798	-10.0	0.042	Equus caballus	(K)GSAHGHSSSGLEK(R)	DOT1-like, histone H3 methyltransferase
XP_001489246	10.2	0.042	Equus caballus	(R)EETNAEmLRQELDRER(Q)	GRAM domain containing 4
XP_001496378	11.5	0.042	Equus caballus	(R)ELHLTEACSSMSFCSPGNEDQNETEGK(E)	Uncharacterized protein C1orf228 homolog
XP_001915320	10.3	0.042	Equus caballus	(R)CIIQMQGNSTSIINPKNPKEAPK(S)	Kinesin family member 1B
XP_001916685	10.2	0.042	Equus caballus	(R)VSDGEQAKmSREISISELECK(D)	Microtubule-actin crosslinking factor 1
XP_001490688	-10.0	0.042	Equus caballus	(K)AVVHGIVLGVVPPPIPEPDGCK(S)	Epididymal secretory protein E1-like
XP_001916798	10.5	0.045	Equus caballus	(R)NIVFQSAVPKVMKV(L)	ADP-ribosylation factor-binding protein
XP_001499311	10.7	0.042	Equus caballus	(R)QLTQHAVEGDQDVR(L)	Alpha-2-HS-glycoprotein-like
XP_001492602	10.1	0.042	Equus caballus	(R)DFHINLFQVLPWLK(E)	Complement factor B
XP_001497655	11.02	0.042	Equus caballus	(R)JTTSRPFLTQKETSPGGSNQEIHKSGLLPLK(A)	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60
XP_001146302	10.3	0.042	Pan troglodyes	(R)VFPEDKtmRYAQmLEK(A)	tRNA-dihydrouridine (16/17) synthase [NAD(P)(+)]like
XP_001488557	-10.0	0.042	Equus caballus	(R)RFGTMSSmSGADDTVYMEYHSSR(S)	TFIIH basal transcription factor complex helicase XPB submit-like
XP_852467	11.4	0.042	Canis familiaris	(R)GEGGAVPGASARGGLTAGRGR(R)	Homeobox D1
XP_001490256	11.0	0.042	Equus caballus	(R)RSSSQPELFEISRQSPESERDIFGAS(-)	Potassium voltage-gated channel, subfamily H (eag-related), member 1
XP_001915324	10.2	0.042	Equus caballus	(K)LHLLPGASKNKLR(T)	Rabphilin-3A-like
XP_001917584	11.9	0.042	Equus caballus	(R)ImAENAAAGISAPSTSPFFYK(A)	Titin
XP_001491191	-10.0	0.042	Equus caballus	(R)GCCFDSSIPR(V)	Trefoil factor 3-like
XP_001495776	10.7	0.042	Equus caballus	(K)VEPPHSHEDLTDDLSTR(S)	Uncharacterized protein KIAA0195-like

Table 2 (Continuation). Identified bronchoalveolar lavage fluid proteins that were differentially expressed (q-value ≤ 0.1 and fold change ≥ 10) between heaves-affected and healthy control horses after a hay exposure trial. A positive fold change indicates over-expression in control horses; a negative fold change indicates over-expression in heaves-affected horses. Bold text indicates proteins previously identified in equine BALF. (6)

RefSeq or GenBank #	Fold		Species	Peptide Sequence	Gene Name
	Change	q-value			
XP_001493758	-10.0	0.042	Equus caballus	(K)KmNHILQWLHPGKTKGK(S)	Protein FAM75D1-like
XP_001917902	10.2	0.042	Equus caballus	(-)mAAPSAVAVASTRLGSHSQGG GLR(L)	Similar to WAS protein
XP_001501063	-10.0	0.042	Equus caballus	(R)ARAAAPNVPCGVCVSAEK(V)	WD repeat-containing protein 1-like
XP_001916340	11.4	0.042	Equus caballus	(K)KVLHAESTKTSLLR(S)	von Willebrand factor A domain containing 3B
NP_001075972	11.5	0.042	Equus caballus	(K)ADFAEVSK(I)	Albumin
NP_001075972	11.6	0.042	Equus caballus	(R)RPCFSALELDEGYVPK(E)	Albumin
NP_612411	-10.0	0.042	Homo sapiens	(R)ELITNKMFESEDSR(N)	SP140 nuclear body protein-like
NP_001074407	11.4	0.042	Rattus norvegicus	(R)EDRDPSTKAHR(N)	Taste receptor, type 2, member 124

horses did develop pulmonary inflammation due to the experimental exposure, as reported in previous studies [20-22].

The proteome of BALF in healthy horses has previously been reported [6]. Although both the current study and previous report used nanospray ionization, the equipment, sample preparation, and MS techniques differed between the studies. Both studies also used different software for protein identification, but the same database (NCBI) was utilized for identifications. Our study identified 250 unique proteins based on tandem MS performed on BALF from two animals, whereas the other study identified 582 unique proteins based on tandem MS performed on two samples of pooled BALF (3 horses in each pool). Forty-three proteins were identified in both studies, including secretoglobin and transferrin. Further, this study was the first to use proteomic techniques to compare protein concentrations between groups of healthy and heaves-affected horses. We were able to find 190 peptides and identify the parent proteins for 33 peptides that are differentially expressed between heaves-affected and control horses. Secretoglobin and transferrin were chosen for validation in this study because previous research has implicated these proteins in inflammatory airway disease in horses and humans [14, 23-30]. These proteins were also selected because equine-specific reagents are available for both secretoglobin and transferrin, whereas equine-specific reagents are not available for many of the other proteins we identified in this experiment.

The lack of significant difference in BALF secretoglobin expression in both the proteomics data and the Western blot data between heaves-affected and control horses after experimental exposure was surprising, as a study from another laboratory found lower secretoglobin expression in heaves-affected horses compared to control horses, both while the heaves-affected horses were in remission from disease and after an experimental exposure trial [14]. Secretoglobin mRNA expression was also found to be down-regulated in horses with summer pasture-associated heaves [23]. Although there was no statistical difference ($p = 0.067$), we did see a tendency for secretoglobin expression to decrease with exposure, both in the heaves-affected and control horses. The same Western blot reagents were used to detect equine secretoglobin in this study as in the study by Katavolos et al [14], so the differences in the results are likely due to factors outside the assay. The horses were exposed for a similar length of time in both studies. Based on comparison of airway neutrophilia, it appears that the experimental exposure trial reported here, while sufficient to induce clinical signs in five heaves-affected horses, did not elicit the same degree of pulmonary inflammation as was seen in the study by Katavolos et al. If the exposure had elicited a stronger pulmonary inflammatory response or been conducted for a longer period of time, we may have seen a statistically significant change in secretoglobin expression. There was a large amount of individual variation in secretoglobin expression, particularly at the exposure measurement. A power calculation showed that a minimum of 15 horses per group

would have been needed in order to achieve statistical significance between baseline and exposure secretoglobulin levels (Type I error rate = 0.05 and power = 0.80).

Asthmatic people and heaves-affected horses have a lower percentage of secretoglobulin positive airway epithelium cells, visualized with immunohistochemistry, compared to controls [14, 26]. A recent study in heaves-affected horses showed that BALF neutrophils have higher intracellular concentrations of secretoglobulin than BALF neutrophils from control horses, and this difference is magnified during exacerbation of clinical signs [27]. The same study also demonstrated that secretoglobulin causes a decrease in oxidative burst in peripheral neutrophils of healthy horses, while also causing an increase in phagocytosis capacity. Secretoglobulin exerts anti-inflammatory effects by inactivating phospholipase A2, decreasing pro-inflammatory cytokine production, and altering phagocyte function, and may also attenuate the inflammatory response to endotoxin [28]. The reason for marked airway inflammation in heaves-affected horses during exacerbation despite increased levels of secretoglobulin within neutrophils is unclear, but it may be because secretoglobulin production by airway epithelium is exhausted and insufficient to keep up with demand, or that secretoglobulin effects on neutrophils in horses with heaves are impaired.

Transferrin expression was significantly higher in the control horses at baseline than in the heaves-affected horses at exposure, and the tendency was for transferrin to decrease with exposure. Again, the pulmonary inflammation seen in the control horses upon hay exposure in this study may have caused decreased transferrin expression, resulting in a lack of statistical difference in transferrin expression between heaves-affected horses and controls. It is also possible that a longer exposure trial, or a more severe exposure, may have resulted in a stronger pulmonary inflammatory response, and a significant difference in transferrin expression between heaves-affected and control horses. Transferrin expression also showed a large amount of individual variation and power calculation showed that a minimum of 7 horses per group would have been needed in order to achieve statistical significance between the secretoglobulin levels in control and heaves-affected horses after exposure to hay (Type I error rate = 0.05 and power = 0.80).

In humans with cystic fibrosis, no significant difference in BALF transferrin concentrations between affected patients and healthy controls has been reported [29]. BALF transferrin was also evaluated in healthy people after exposure to swine dust in a confinement operation for three hours. Within 24 hours from the start of exposure to swine dust, BALF transferrin concentrations increased by 2.6 fold. The change in transferrin concentration was correlated with the total dust exposure levels [30]. The trend for a decrease in BALF transferrin expression in heaves-affected horses after the experimental exposure trial seems to be in contrast to this study. In the study presented here, samples were not taken from the heaves-affected horses in the experimental exposure until their clinical signs and lung function reached

a certain level of severity. It is possible that if these horses had been tested earlier in the course of disease (more acute inflammation) or later in the exposure (chronic inflammation), transferrin expression may have been different.

We also found that secretoglobulin and transferrin expression were stable over time in heaves-affected horses that were in clinical remission from the disease and in healthy controls. The stability of BALF secretoglobulin and transferrin may make these proteins potential targets for monitoring subclinical inflammation, or response to environmental changes.

We performed proteomics on cell-free BALF supernatant at one time point after experimental exposure. Different protein expression levels may have been obtained by evaluating the cellular fractions of BALF rather than the cell-free supernatant, as secretoglobulin has been found to be concentrated in BALF neutrophils in horses [27]. In addition, if repeated measurements had been performed during the exposure trial, we would have been able to track the changes in secretoglobulin and transferrin secretion. As clinical signs of heaves tend to worsen with prolonged environmental exposure, it is likely that we would have found altered expression patterns for both secretoglobulin and transferrin over time.

5. Concluding Remarks

In this study, we demonstrated that proteomic techniques can be used for discovery of proteins in BALF of horses with heaves. Two proteins identified using mass spectrometry were confirmed using Western blot. Many additional proteins identified with proteomics show differential expression between the control and heaves-affected horses. Since only one pulmonary disease was evaluated in this study, it is likely that just a fraction of the differentially expressed proteins identified here are specific biomarkers for heaves and that some of the differentially expressed proteins may be non-specific indicators of pulmonary inflammation (e.g., immunoglobulins). Further validation, including confirmation of the proteins' identity using techniques such as Western blot, and investigation of protein expression in other pulmonary diseases must occur before a protein can be considered a biomarker for heaves.

^a Equine Senior, Purina Mills, LLC, Gray Summit, MO; ^b Dormosedan, SmithKline Beecham Animal Health, West Chester, PA; ^c Torbugesic, Fort Dodge Laboratories, Fort Dodge, IA; ^d Olympus, Center Valley, PA; ^e Thermo Scientific Pierce BCA Protein Assay Kit, Thermo Fisher Scientific, Inc., Rockford, IL; ^f 1100 Series LC equipped with HPLC Chip interface, Agilent, Santa Clara, CA; ^g Agilent, Santa Clara, CA; ^h Available through special arrangement with Purdue University, West Lafayette, IN; <http://proteobbc.purdue.edu:7080/pipeline/>; ⁱ National Institutes of Health, Bethesda, MD; ^j Invitrogen, Carlsbad, CA; ^k Pall Life Sciences, Ann Arbor, MI; ^l Gift from Dr. Dorothee Bienzle, Ontario Veterinary College, University of Guelph,

Canada; ^m Jackson ImmunoResearch Laboratories, Inc, West Grove, PA; ⁿ R&D Systems, Minneapolis, MN; ^o Bethyl Laboratories, Inc., Montgomery, TX; ^p Statistica, StatSoft, Inc, Tulsa, OK; ^q MATLAB, The Mathworks, Inc, Natick, MA

6. Supplementary material

Supplementary data and information is available at: <http://www.jiomics.com/index.php/jio/rt/suppFiles/112/0>

Supplementary material includes:

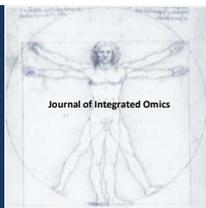
- *Supplementary file 1*: Supplementary information for Materials and Methods
- *Supplementary file 2*: Peptide identification of heaves-affected and control horses
- *Supplementary file 3*: The peptide sequence of the equine secretoglobin precursor protein (GenBank Accession #AAW83220), including an identification of peptides that were found with LC-MS/MS analysis
- *Supplementary file 4*: The peptide sequence of the equine serotransferrin precursor protein (GenBank Accession #NP_001075415), including an identification of peptides that were found with LC-MS/MS proteomic analysis
- *Supplementary file 5*: Spearman rank correlations between the density of secretoglobin or transferrin on Western blot and BALF cytology parameters

Acknowledgements

This work was completed at Purdue University, West Lafayette, Indiana, and was supported by the state of Indiana, the Purdue University College of Veterinary Medicine Research account funded by the total wager tax, and the American College of Veterinary Internal Medicine Foundation. We thank Dr. Dorothee Bienzle (University of Guelph, Guelph, Ontario, Canada) for her generous donation of equine CCSP peptide, recombinant equine CCSP and anti-equine CCSP serum. Special thanks to Donna Griffey, RVT, Dr. Katy Ivester, Dr. Kristi Goncarovs, Dr. Jordan Hammer, and Kathy Anderson for their technical help.

References

1. V. Bracher, R. von Fellenberg, C.N. Winder, G. Gruenig, M. Hermann, A. Kraehenman, *Equine Vet. J.* 23 (1991) 136-141.
2. J.W. Hotchkiss, S. W. Reid, R.M. Christley, *Equine Vet. J.* 39 (2007) 301-308.
3. United States Department of Agriculture National Agricultural Statistics Service, Indiana Equine Summary. http://www.nass.usda.gov/Statistics_by_State/Indiana/Equine_Summary/index.asp
4. M. Leclere, A. Lavoie-Lamoureux, J.P. Lavoie, *Respirology.* 16 (2011) 1027-1046.
5. Ramseyer, C. Gaillard, D. Burger, R. Straub, U. Jost, C. Boog, E. Marti, V. Gerber, *J. Vet. Intern. Med.* 21 (2007) 149-156.
6. L.A. Bright, N. Mujahid, B. Nanduri, F.M. McCarthy, L.R.R. Costa, S.C. Burgess, C.E. Swiderski, *Animal Genetics.* 42 (2011) 395-405.
7. J. Hirsch, K.C. Hansen, A.L. Burlingame, M.A. Matthay, *Am. J. Physiol. Lung Cell Mol. Physiol.* 287 (2004) L1-23.
8. X. Zhang, A. Fang, C.P. Riley, M. Wang, F.E. Regnier, C. Buck, *Anal. Chim. Acta.* 664 (2010) 101-113.
9. U. Rajcevic, S.P. Niclou, C.R. Jimenez, *Front. Biosci.* 14 (2009) 3292-3303.
10. N.E. Robinson, *Equine Vet. J.* 33 (2001) 5-19.
11. D.B. Tesarowski, L. Viel, W.N. McDonell, *Am. J. Vet. Res.* 57 (1996) 1214-1219.
12. L.L. Couetil, F.S. Rosenthal, C.M. Simpson, *J. Appl. Physiol.* 88 (2000) 1870-1879.
13. X. Zhang, J.M. Asara, J. Adamec, M. Ouzzani, A.K. Elmagarmid, *Bioinformatics.* 21 (2005) 4054-4059.
14. P. Katavolos, C.A. Ackerley, L. Viel, M.E. Clark, X. Wen, D. Bienzle, *Vet. Pathol.* 46 (2009) 604-613.
15. L.L. Couetil, M.P. Ward, *J Am. Vet. Med. Assoc.* 223 (2003) 1645-1650.
16. F.R. Bertin, K.M. Ivester, L.L. Couetil, *Equine Vet. J.* 43 (2011) 393-398.
17. C.L. Reyner, B. Wagner, J.C. Young, D.M. Ainsworth, *Am. J. Vet. Res.* 70 (2009) 1277-1283.
18. M.E. Cordeau, P. Joubert, O. Dewachi, Q. Hamid, J.P. Lavoie, *Vet. Immunol. Immunopathol.* 97 (2004) 87-96.
19. N.E. Robinson, C. Berney, A. Behan, F.J. Derksen, *J. Vet. Intern. Med.* 23 (2009) 1247-1253.
20. G.M. Tremblay, C. Ferland, J.M. Lapointe, A. Vrins, J.P. Lavoie, Y. Cormier, *Equine Vet. J.* 25 (1993) 194-197.
21. C. Kleiber, B.C. McGorum, D.W. Horohov, R.S. Pirie, A. Zurbriggen, R. Straub, *Vet. Immunol. Immunopathol.* 104 (2005) 91-97.
22. Desjardins, C. Theoret, P. Joubert, B. Wagner, J.P. Lavoie, *Vet. Immunol. Immunopathol.* 101 (2004) 133-141.
23. C.S. Venugopal, L.C. Mendes, J.R. Peiro, S.S. Laborde, A.M. Stokes, R.M. Moore, *Am. J. Vet. Res.* 71 (2010) 476-482.
24. N. Shijubo, Y. Itoh, T. Yamaguchi, F. Sugaya, M. Hirasawa, T. Yamada, T. Kawai, S. Abe, *Lung.* 177 (1999) 45-52.
25. Q. Ye, M. Fujita, H. Ouchi, I. Inoshima, T. Maeyama, K. Kuwano, Y. Horiuchi, N. Hara, Y. Nakanishi, *Respiration.* 71 (2004) 505-510.
26. N. Shijubo, Y. Itoh, T. Yamaguchi, A. Imada, M. Hirasawa, T. Yamada, T. Kawai, S. Abe, *Am. J. Respir. Crit. Care Med.* 160 (1999) 930-933.
27. P. Katavolos, C.A. Ackerley, M.E. Clark, D. Bienzle, *Vet. Immunol. Immunopathol.* 139 (2011) 1-9.
28. J.C. Snyder, S.D. Reynolds, J.W. Hollingsworth, Z. Li, N. Kaminski, B.R. Stripp, *Am. J. Respir. Cell. Mol. Biol.* 42 (2010) 161-171.
29. S.W. Stites, M.W. Plautz, K. Bailey, A.R. O'Brien-Ladner, L.J. Wesselius, *Am. J. Respir. Crit. Care Med.* 160 (1999) 796-801.
30. O. Vesterberg, L. Palmberg, K. Larsson, *Int. Arch. Occup. Environ. Health.* 74 (2001) 249-254.



JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM



ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.114

Electrophysiological and Proteomic Studies of *Protobothrops mangshensis* Venom Revealed Its High Bioactivities and Toxicities

Yu Liu^{1,2,#}, Xiaojuan Wang^{2,#}, Zhe Wu¹, Ping Chen², Yang Xiang¹, Dongsong Nie¹, Jianmin Yi¹, Jizu Yi^{1,*}, and Zhonghua Liu^{2*}

¹College of Chemistry and Chemical Engineering, Hunan Institute of Science and Technology, Yueyang, 414006, China; ²College of Life Sciences, Hunan Normal University, Changsha, 410081, China; [#]Those authors contributed equally to this study.

Received: 28 November 2012 Accepted: 07 December 2012 Available Online: 20 December 2012

ABSTRACT

Snake venoms contain rich components having medical and biotechnological values. The proteomic characterization of snake venom proteome has thus potential benefits for basic research, clinical diagnosis, and development of new drugs for a variety of diseases. *Protobothrops mangshensis* is a monotypic genus of pit viper known only from Mountain Mang in Hunan Province of China, and represents the largest and the most spectacular snake among Asian venomous snakes. The venom of *Protobothrops mangshensis* exhibits a high coagulant activity on bovine and human fibrinogen and human plasma, a high phosphodiesterase activity and an arginine ester hydrolytic activity. In this study, the *Protobothrops mangshensis* venom was analyzed by 2D-gel electrophoresis separation, subsequently by in-gel digestion, MS/MS identification, and enzymatic activity analysis. Our results demonstrated that *Protobothrops mangshensis* venom comprised highly functional proteins and/or enzymes, and each of these proteins displayed multiple isoforms separated in 2-DE. Approximately 59.4% of the identified total 143 proteins had enzymatic activities and 24.5% were involved ion channels, representing highly complex and extensive bioactivities of the snake venom. The identified toxins included six protein families: serine proteinases, L-amino acid oxidase, phospholipases A2, C-type lectin-like proteins, cysteine-rich secretory proteins and metalloproteinase-disintegrin, and were correlated well with the clinical manifestations by *Protobothrops mangshensis* bite such as coagulopathy, oedema, hypotensive and tissue damaging effects. Electrophysiological studies showed that the snake venom inhibited tetrodotoxin-resistant (TTX-R) Na⁺ currents. All our results in this study provided the first functional proteomics of *Protobothrops mangshensis* venom.

Keywords: Snake venom; Venomics; Protein family; 2-D electrophoresis; Proteomics.

Introduction

A snake venom fluid contains varieties of proteins and peptides with different biological activities including enzymatic, cytotoxic, neurotoxic, and hemagglutinic activities [1]. Some of the components found in a snake venom sample are of interests as tools for studying neurophysiology and/or as potential lead structures for new drug development. Recently there were several reports that snake venoms were used in drug development for novel human disease therapeutics. These included the antihypertensive drug captopril, the molecule modeled from the venom of the Brazilian arrowhead viper (*Bothrops jararaca*) [2]; the anticoagu-

lant Integrilin (eptifibatide), one heptapeptide derived from a protein found in the venom of the American southeastern pygmy rattlesnake (*Sistrurus miliarius barbouri*) [3]; An-crod, one compound isolated from the venom of the Malaysian pit viper (*Agkistrodon rhodostoma*) for treating heparin-induced thrombocytopenia and stroke [4]; and alfineprase, one novel fibrinolytic metalloproteinase for thrombolysis derived from southern copperhead snake (*Agkistrodon contortrix*) venom [5]. Due to the limits of the sensitivity and resolution of commonly used methods for protein separation, the complexity and diversity of components in snake

*Corresponding author: Dr. Jizu Yi, College of Chemistry and Chemical Engineering, Hunan Institute of Science and Technology, Yueyang, 414006, China, phone number: +86-730-8640122, fax: +86-730-8640122, e-mail address: Jizu_Yi@Yahoo.com; Dr. Zhonghua Liu, College of Life Sciences, Hunan Normal University, Changsha, 410081, China, phone number: +86-731-88872556, fax: +86-731-88861304, e-mail address: liuzh@hunnu.edu.cn

venom were usually underestimated. Many compounds with low abundance were often neglected. Two-dimensional polyacrylamide gel electrophoresis (2-DE) was one of efficient methods for separating several hundred proteins based on the differences in their pIs and molecular masses. The immobilized pH gradient (in the first-dimensional) gel strip-based 2-DE separation, followed by image analysis, in-gel digestion, digested protein sample handling, and mass spectrometry (MS) analysis together with bioinformatic algorithms for searching sequences in databases, present currently a typical proteomic method well suitable for global and detailed studies of the entire proteome from a certain cell type, or a tissue of an organism, or a body fluid such as a snake venom. The progress in MS technology has dramatically accelerated the application of proteomics during recent years. Venoms from various venomous animals have been analyzed using MS technologies such as MALDI-TOF and LC/ESI-Q-TOF [6–10]. These studies have demonstrated the advantages of using proteomic approaches in venom mapping [8–10], therapeutic target screening, and species' classification [11]. *Protobothrops mangshensis* was firstly discovered in the forests on Mount Mang at elevations from 700 to 1300 m, located at Pingkeng District, Yizhang County, Hunan Province of China [12, 13]. To date, the known place where these snakes live was restricted to this mountain area ranging only a few tens of square kilometers [12, 13]. An adult *Protobothrops mangshensis* grows to body length of about 2.0 m, weight of about 2–4 kg, and is known to prey mainly on rodents [12]. It was also named as *Trimeresurus mangshanensis*, *Zhaoermia mangshanensis* or *Ermia mangshanensis* [14, 15], but recently reclassified as the genus *Protobothrops* based on its mitochondria gene analysis [16].

The crude venom from a *Protobothrops mangshensis* contained a wide variety of proteins and peptides, and was highly lethal to mice with an intraperitoneal LD₅₀ of 4.2 mg/kg body weight [17]. The envenoming elicited coagulopathy, local hemorrhage, inflammation, edema, blood blisters, myonecrosis, and severe pain on human victims [17]. The treatment of an envenomation by a snake bite was critically dependent on the availability of effective antivenoms. Detailed knowledge of the identities and relative amounts of the different toxins in the given venom was thus required for generating immunization protocols to elicit toxin-specific antibodies showing greater specificity and effectiveness than conventional methods relying on the immunization of large mammals with whole venom. Here, we reported an electrophysiological and proteomic study of the venom collected from *Protobothrops mangshensis*. Our results demonstrated that the venom contained a variety of enzymes and proteins including serine proteinases, L-amino acid oxidase, phospholipases A2, C-type lectin-like proteins, cysteine-rich secretory proteins, metalloproteinase-disintegrin, and more others. The venom toxin composition was correlated with the clinical manifestation of the *Protobothrops mangshensis* bite, explaining its pathological effects such as coagulopathy, oedema, as well as hypotensive and tissue damaging effects.

Electrophysiological results showed that the venom fractions inhibited tetrodotoxin-resistant (TTX-R) Na⁺ currents, but not on tetrodotoxin-sensitive (TTX-S) Na⁺ currents. Together, these results provided a detailed characterization of the venom proteome of the snake for a deeper understanding the biological effects of the venom, and served as a starting point for further structure-function studies of individual toxins.

1. Materials and methods

1.1 Chemicals

Sephadex G-75, IPG (pH 3–10), DryStrips (3–10 linear), cover fluid, agarose, and colloidal Coomassie Brilliant Blue (CBB) were purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). Dithiothreitol (DTT), iodoacetamide, trypsin, and trifluoroacetic acid (TFA) were obtained from Sigma (St. Louis, MO, USA). Acrylamide, bis-acrylamide, urea, thiourea, glycine, tris, α -Cyano-4-hydroxycinnamic acid (CHCA), and sodium dodecyl sulfate (SDS) were from of Amresco (Solon, OH, USA). Acetonitrile (ACN, chromatogram grade) and other chemicals (analytical grade) are from Sinopharm Group (Shanghai, China). Deionized water prepared with an Aquapro high-end water treatment solution provider system (Ever Young, China) was used for all buffers in this study.

1.2 Snake venom

Protobothrops mangshensis venom was collected from two adult female snakes maintained in Mangshan Institute of Snakes. The snakes were manually restrained and venom fluids were collected with an opened glass container. Approximately 2.0 ml of yellowish collected venom liquid was lyophilized immediately at cold temperature to minimize possible preanalytical variability during the sample handling [18, 19]. Total ~ 1200 mg dry powder of the venom sample was obtained and stored at 4°C until usage.

1.3 Electrophysiological studies

Whole cell sodium currents were recorded from rat dorsal root ganglion (DRG) cells at room temperature. Recording pipettes were made from borosilicate glass capillary tubing, and their resistances were 1–2 megaohms when filled with internal solution containing: 135.0 mM CsF, 10.0 mM NaCl in 5.0 mM HEPES at pH 7.0. External bath composition was: 30 mM NaCl, 5.0 mM CsF, 25 mM D-glucose, 1.0 mM MgCl₂, 1.8 mM CaCl₂, 20.0 mM triethanolamine-chloride, and 70 mM tetramethylammonium in 5.0 mM HEPES (pH 7.4). Ionic currents were filtered at 10 kHz and sampled at 3 kHz on EPC-9 patch clamp amplifier (HEKA Electronics). Linear capacitive and leakage currents were subtracted by using a P/4 protocol. Experimental data were acquired and analyzed by the software program pulse+ pulsefit 8.0 (HEKA

Electronics, Germany). The toxin dissolved in external solution until desired concentrations was applied onto the surface of experimental cells by low pressure injection with a microinjector (IM-5, Narishige, Japan).

1.4 Separation of venom proteins by two-dimensional gel electrophoresis (2-DE)

2-DE was performed according to the method described previously [20]. IEF was carried out on an IPGphor system (Amersham Pharmacia Biotech, USA). Venom protein samples (300 mg) pooled from protein fractions collected after gel-filtration chromatography were first mixed with a rehydration solution containing 8 M urea, 2 M thiourea, 4.0 % CHAPS, 20 mM Tris-base, 18.0 mM DTT, and traces of bromophenol blue in 0.5 % (v/v) IPG buffer to give a total volume of 350 mL, and applied to IPG dry strips (pH 3-10, 180×30×0.5 mm) for the rehydration up to 14 hrs. IEF was then carried out at step-n-hold conditions: 500 V for 1 h; 1000 V for 1 h; and 8000 V for 6 hrs. at 50 Ma/strip. After focusing, the strips were soaked for 20 min in reduction solution (6 M urea, 30 % glycerol, 2 % SDS, and 125 mM DTT) followed by 20 min in alkylation solution (6 M urea, 30 % glycerol, 2.0 % SDS, and 125 mM iodoacetamide). The second dimensional separation with SDS-PAGE based on the molecular size differences was performed in 10 % polyacrylamide gels using a Protean II system (Bio-Rad, Hercules, CA, USA). The gels were then stained with CBB G250 and scanned using ProXPRESS 2D Proteomic Imaging System (Perkin-Elmer, Waltham, MA, USA). Protein spots were detected and analyzed using PDQuest software Version 6.1 (Bio-Rad).

1.5 In-gel protein digestion and MS analyses using ESI-Q-TOF and MALDI-TOF-TOF MS

The CBB-stained protein spots were excised and digested in-gel with trypsin as described previously [21]. Peptide mixtures from in-gel digestion were extracted and analyzed by an ESI-Q-TOF MS/MS (Waters, Milford, MA, USA) with the nanoelectrospray for ionization as described previously [22]. We selected the candidate peptides for identifications based on the probability-based Mowse scores (total score) which exceeded their thresholds indicating a significant (or extensive) homology ($p < 0.05$), and referred them as “hits” as defined by Matrix Science, Ltd. (London, UK). Proteins identified with two or more peptides matched with a Mowse score greater than 40 of each were validated without further investigations. Those with at least two peptides matched with a score less than 40 and greater than 20 of each peptide were systematically checked manually to confirm or cancel the MASCOT suggestion. For a protein identified by only one matched peptide, the Mowse score of the MS/MS search must exceed 40, and the peptide sequence was checked manually to confirm the protein ID.

Tandem mass spectra acquired on an ESI-Q-TOF mass

spectrometer were interpreted de novo using masslynx 4.0 software (Waters). Contiguous stretches of seven or more amino acid residues with a 100% confidence call using the software’s default parameters were collected and matched to the NCBI non-redundant protein database using the protein BLAST algorithm (<http://www.ncbi.nlm.nih.gov/BLAST/>).

The tryptic digestion peptide mixture from a 2-DE gel spot without a matched result by ESI-Q-TOF MS/MS were loaded onto an AnchorChip target plate (Bruker-Daltonics, Germany) according to previous report [23]. Peptide identification was performed using a MALDI-TOF-TOF mass spectrometer (UltraFlex I, Bruker-Daltonics, Germany) equipped with a nitrogen laser (337 nm) and operated in reflector/delay extraction mode for both peptide mass fingerprint (PMF) and MALDI-TOF-TOF MS methods. The MS spectra were obtained by a fully automated mode through the flexControl software (Bruker-Daltonics). An accelerating voltage of 25 Kv was used for PMF. The peaks with S/N ≥ 5 and resolution ≥ 2500 were selected for further TOF-TOF MS analysis from the same target. The acquired PMF MS and peptide TOF-TOF MS spectra were combined and searched using Mascot (Matrix Science) with peptide mass tolerance of 50 ppm, and MS/MS tolerance of 1.0 Da. The protein identifications were accepted when the total Mowse score of the “hit” exceeded the threshold significance score more than 70 ($p < 0.05$).

2. Results and Discussion

2.1 Effects of *Protobothrops mangshensis* venom on ion channel currents

It was widely accepted that voltage-gated sodium channels (VGSCs), voltage-gated potassium channels (VGPCs), and voltage-gated calcium channels (VGCCs) existed in rat dorsal root ganglia (DRG) neurons. By whole-cell patch-clamp recording, we first investigated the effects of the venom from *Protobothrops mangshensis* on voltage-gated Na⁺, K⁺ and Ca²⁺ channels in adult rat DRG neurons. We observed that *Protobothrops mangshensis* venom had no evident effect on tetrodotoxin-sensitive (TTX-S) Na⁺ currents (Fig. 1A). However, it was able to inhibit tetrodotoxin-resistant (TTX-R) Na⁺ currents and the inhibition was dose-dependent (Fig. 1B). The venom had also no evident effect on delayed rectifier potassium current (Fig. 1C) and low voltage-activated calcium current (Fig. 1D), but a little effect on high voltage-activated calcium current (Fig. 1E). The results indicated that the venom components blocked effectively on TTX-R sodium channels and less effectively on high voltage-activated calcium channels. To further confirm this effect, we separated the venom by a gel-filtration chromatography and collected the fractions (I, II, III and IV) for their individual tests (Fig. 2). We observed that fractions I and II had no significant effect on TTX-R sodium currents (Figs. 2B & 2C) and the inhibitory effects of fractions III and IV increased on TTX-R sodium currents (Figs. 2D & 2E).

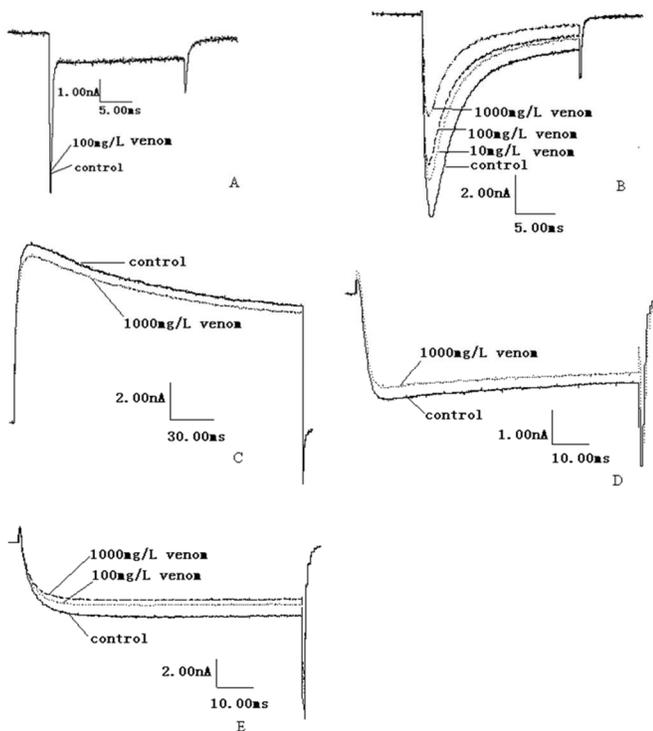


Fig. 1. Effects of *Protobothrops mangshanensis* crude venom on voltage-gated ion channels on rat DRG cells. (A) The TTX-S Na^+ currents were elicited by 50 ms voltage steps to -10 mV from a V_h of -80 mV. The currents were almost unaffected by 100 mg/L venom ($n=4$); (B) TTX-R sodium currents were elicited by depolarizing the cell from a holding potential of -80 mV to -10 mV. The duration of the test pulses was 50 ms. 10 mg/L venom caused a significant decrease in peak current amplitude. The inhibition was dose-dependent. (C) The delayed rectifier K^+ currents were elicited by a 30 ms depolarization to +30 mV from a V_h of -80 mV. The currents were almost unaffected by 1,000 mg/L venom. (D) Low voltage-activated currents were induced by a 150 ms depolarizing potential of -30 mV from a holding potential -90 mV, and current traces were not changed before and after the application of 1,000 mg/L venom. (E) High voltage-activated currents were elicited by 150 ms depolarizing voltage of 0 mV from a holding potential of 40 mV, and current traces were slightly changed before and after the application of 1,000 mg/L venom.

The results demonstrated that the TTX-R sodium channel blockers presented within the composition of the fractions III and IV.

2.2 Proteomic profile of venom proteins of *Protobothrops mangshanensis*

Due to the limits in the detection sensitivity and resolution of commonly used methods for protein separation, the complexity and diversity of snake venom proteins were traditionally underestimated. Many protein components with low abundance were often omitted in previous investigations. Here, we used a proteomic strategy to obtain more comprehensive identifications of the snake venomome (or venom proteome). The venom was first separated by 2-D gel electrophoresis (Fig. 3), followed by in-gel trypsin digestion, and MS as well as MS/MS analyses.

Due to the absence of a snake venom database, the LC MS/

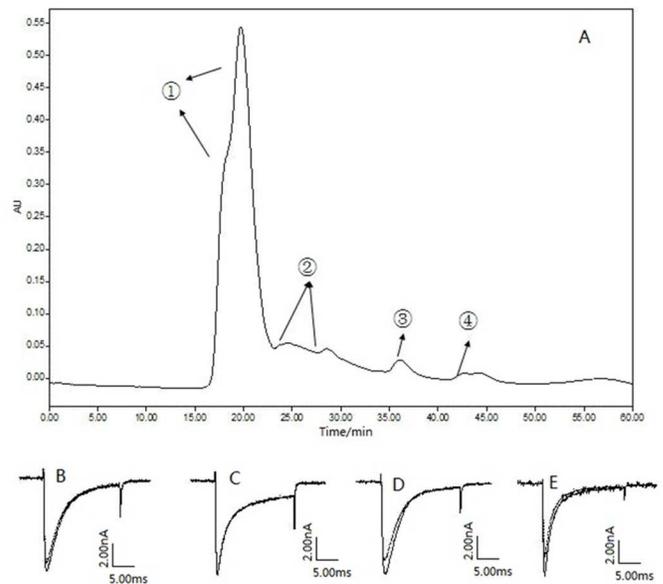


Fig. 2. Effects of *Protobothrops mangshanensis* venom fractions after Gel filtration chromatography on TTX-R sodium currents. The current traces were not changed before and after the application of fractions ① (B) and ② (C), but decrease moderately after adding fraction ③ (D) and fraction ④ (E). The results showed that the TTX-R sodium channel blockers present in these two fractions.

MS and MALDI TOF-TOF MS spectra were searched against the theoretical peptide MS of Metazoa proteins in NCBI and Swiss-Prot databases, with total 43 proteins being identified from Swiss-Prot. Some spots from 2-DE displayed high-quality ESI-Q-TOF MS/MS spectra but resulted in no matched sequences. For those spots the spectra were subjected to BLAST from NCBI, and 100 more proteins were identified. We categorized the identified proteins according to their functions on the basis of universal GO annotation terms, and found that 59.4 % of the total proteins were enzymes including types of proteinases, oxidases, phospholipase, and so on; 24.5 % were ion channel proteins; 9.8 % behaved as cellular binding and structural proteins; 3.5 % were signal transport proteins, and only 2.8 % other proteins were unknown in their functions. The fact that the greatest percentage (~60 %) of the identified proteins was enzymes and the second (~25 %) were associated with ion channels indicated that the biological activities of a venom fluid were extremely intensive and extensive.

2.3. Variability, bioactivity and toxicity

The identified toxins were found mainly belonged to the following six protein families: phospholipases A2, L-amino acid oxidases (LAAOs), C-type lectins, and cysteine-rich secretory proteins (CRISP), metalloproteinase-disintegrin, and serine proteases (see Supplementary material, Table 1). Specifically, among the identified venom serine proteinases (VSP) included VSP 1 (spots # 61, 62, 63, 64, 70, 97, 98, 101), VSP 2 (spot 194), VSP 2A (spot 188), and VSP 2C (spots 67, 68, 78, 79) (Fig. 3 and Supplementary material, Table 1) in

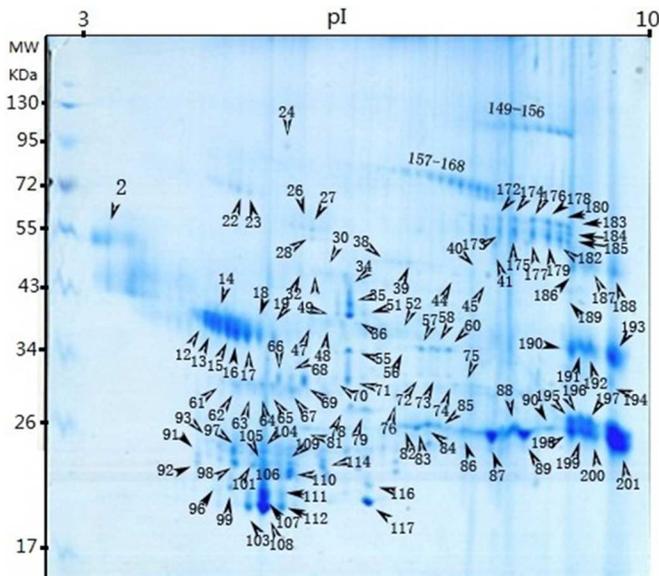


Fig. 3. 2-DE map of proteins from *Photobothrops mangshanensis* venom. Protein spots were visualized by CBB G 250 staining. The spots analyzed by MS were designated by numbers; their identifications were listed in Table 1.

addition to several other thrombin-like serine proteinases (see below), implying its high proteolytic activities of the snake venom fluid.

Coagulopathy was one of the major toxic effects exerted by the *photobothrops mangshanensis*. The venom proteins revealed by our proteomic analysis were in accordance with this clinically observed blood clotting disorders caused by a snake bite. Indeed, several thrombin-like proteases associated with blood clotting cascade were identified, among which were pallabin (spot 56), flavoxobin (spot 65), mucrofibrase-5 (spots 67, 75, 81), halystase-like proteinase (spots 191, 193) and conortrixobin (Fig. 3 and Supplementary material, Table 1). Furthermore, multiple isoforms of the thrombin-like serine proteases were separated by 2-DE and many of these toxins were found in the middle part of the gel (Fig. 3). As one representative of such thrombin-like serine proteinases, conortrixobin was isolated and identified from multiple spots (# 63, 65, 66, 67, 68, and 70 on Fig. 3). The multiple isoforms of a single protein were observed commonly in this study, demonstrating the high variability of the *Protobothrops mangshensis* venom proteins.

In terms of its structure and functions, conortrixobin had strong homology with snake venom serine proteases acting on either fibrinogen or other blood coagulation components. The interaction of conortrixobin with chromogenic substrates indicated a higher specificity for arginine than lysine in the primary subsite and a faster attack to ester than amides [24]. Interestingly, conortrixobin preferentially released fibrinopeptide B from human fibrinogen and activated blood coagulation Factors V and XIII with a rate 250-500-fold lower than human alpha-thrombin. However conortrixobin did not induce the aggregation of thrombocytes, the increase of intracytoplasmatic calcium ions in platelets, and the activation of Factor VIII [24].

However, halystase-like proteinases (spots 191, 193 on Fig. 3) were expected to have the activity similar to that of other thrombin-like snake venom serine proteases: to hydrolyze the fibrinogen chains at the sites different from thrombin and kininogen producing bradykinin, which led to hypotension. No coagulation effect on the human plasma by the identified halystase-like proteases was expected [25].

The coagulation factor IX/factor X-binding proteins were also isolated and identified (Fig. 3, spots 103, 104, 106, 107, 108, 109, and 111). The coagulation factor IX/factor X-binding protein (IX/X-bp) from the venom of *Trimeresurus flavoviridis* was reported to be a heterogeneous two-chain protein, and the structure of each chain was similar to that of the carbohydrate-recognition domain of C-type lectins. The analysis of the binding properties of IX/X-bp revealed that it bound to the gamma-carboxyglutamic acid (Gla)-containing domains of factors IX and X [26]. Here, we also identified galactose-binding lectins (Fig. 3, spots 82, 83, 84, 85, and 86), which recognized the specific carbohydrate structures and agglutinated a variety of animal cells by binding to cell-surface glycoproteins and glycolipids. A galactose-binding lectin was a calcium-dependent protein which showed a high hemagglutinating activity [27].

We also identified zinc metalloproteinase-disintegrin jerdonitin from the photobothrops mangshanensis venom (Fig. 3, spots 44, 49, 52, 58, 73, 74) (also see Supplementary material, Table 1). This protease inhibited ADP-induced human platelet aggregations and belonged to the class II of snake venom metalloproteinases (SVMPs) (P-II class). Different from other P-II class SVMPs, the metalloproteinase and disintegrin domains of this natural protein were not separated. It has two additional cysteine residues located separately in the spacer domain and disintegrin domain. The cysteine residues probably formed a disulfide bond which bound the metalloproteinase and disintegrin domains together during posttranslational processes [28].

L-amino acid oxidase (LAAO) was another major component of the *Photobothrops mangshanensis* venom. Eleven isoforms of this enzyme were detected after 2-DE and MS/MS analysis of tryptic peptides (Fig. 3 and Supplementary material, Table 1). The proteins isolated from the spots 172, 174, 175, 176, 177, 178, 180, 181, 183, 184 and 185 had closely same molecular masses of approximately 58 kDa, a typical mass of LAAO, but different isoelectric points in the alkaline region. Although the knowledge about the biological roles of LAAOs was limited, it was reported that these enzymes induced apoptosis, hemorrhagic effect and cytotoxicity [29]. These activities could be attributed to the high number of LAAO isoforms in the venom. With the similar molecular weight, these multiple isoforms of the LAAO, as well as other protein isoforms, were generated by either post-translation modifications or *in vivo* mutual modifications by the intrinsic venom enzymes. This high protein/enzyme isoforms of the snake venom observed also with many other protein spots in 2-DE (Fig. 3) demonstrated not only the high biological variability but also the complex and multiple

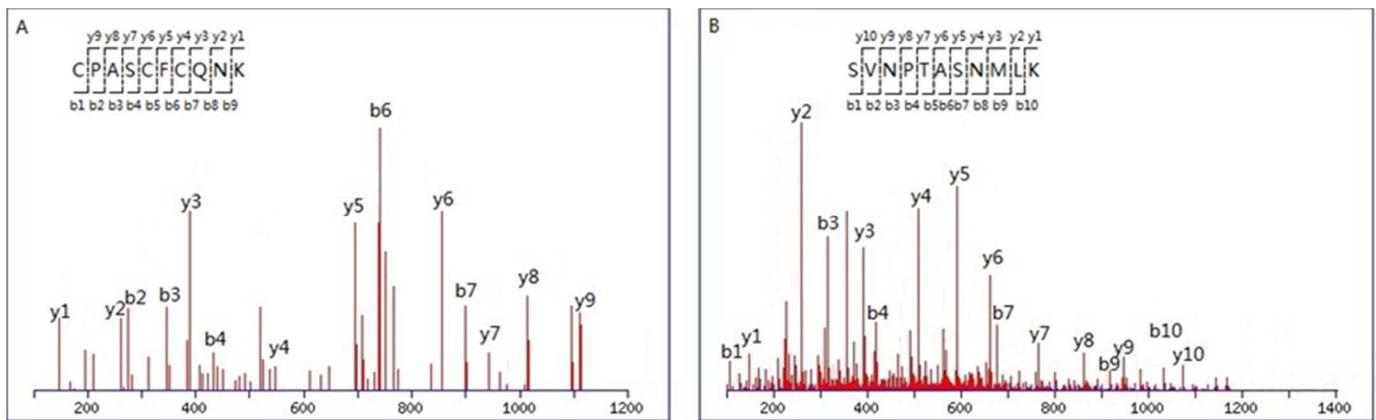


Fig. 4. Typical tandem ESI-MS spectra of two trypsin-digested peptides of Triflin from *Protobothrops mangshanensis* venom. The identification of the Triflin (spot # 57 in Fig. 3) was confirmed with acquired fragment spectra of two peptides: (A) CPASCFCQNK (FM:1270.39) from 636.20 (M/z) and (B) SVNPTASNMLK (FM: 1160.57) from 581.29 (M/z), with both peptides doubly charged. Only corresponding y- and b-ion were indicated.

functionalities of the venom proteins.

Phospholipase A₂s (PLA₂s), as another example, were also multifunctional enzymes that exerted multiple pharmacological activities: presynaptic/postsynaptic neurotoxicity, myotoxicity, cardiotoxicity, antiplatelet, convulsant, hypotensive and oedema-inducing effects [30]. Seven PLA₂s were found in the pooled *Protobothrops mangshanensis*, including an acidic E6-PLA₂, another acidic R6E7-PLA₂, two basic N6-PLA₂s, two basic R6-PLA₂s, and one R49-PLA₂ homolog in the previous report [31]. However only R49-PLA₂ without enzymatic activities was isolated and characterized from an individual male *Protobothrops mangshanensis* specimen [32]. In this study, several PLA₂ components were also identified (Spots 195, 197, 198, 199, 200 in Fig. 3 and see Supplementary material, Table 1). We also did not detect any PLA₂ enzymatic activity in this *Photobothrops mangshanensis* venom sample. The amino acid sequences from *Photobothrops mangshanensis* toxin and other R49-PLA₂ or K49-PLA₂ homologs were aligned with each of 121 or 122 a.a. residues including seven conserved cysteine residues which could cross-linked the proteins by inter or intra molecular disulfide bonds. With the sequence identity of >80%, they were highly expected structurally similar to each other. Since Asp48 was responsible for direct binding to the active site calcium ion, the PLA₂s with basic substitutions at this position could not bind calcium ion which was necessary for its hydrolytic activities. Notably, the R49-PLA₂s were found only in *Protobothrops* venoms, and thus could serve as venom markers of this genus. Furthermore, the R49-PLA₂s were functionally identical to the K49-PLA₂ subfamily. R49-PLA₂ from *Photobothrops mangshanensis* could bind membrane and induced edema and myonecrosis in mice [32], thus contributed significantly to the symptom elicited by *Photobothrops mangshanensis* envenoming. Given the fact that R49-PLA₂s are especially rich in *Photobothrops mangshanensis* venom, certain anti-inflammatory drugs can be potential therapeutic agents in treating the snake bite.

Among many other identified proteins, Triflin, as one of

cysteine-rich secretory proteins (CRISPs), was also detected and identified from multiple spots: 57, 86, 87, 88, 89 and 90 (Figs. 3 & 4). CRISPs were found in the venoms of a wide variety of snake species [33], such as ablomin from Japanese Mamushi snake (*Gloydius blomhoffi*, formerly *Agkistrodon blomhoffi*) [34], latisemin from Erabu sea snake (*Laticauda semifasciata*) [33], ophanin from King Cobra (*Ophiophagus hannah*) [35], piscivorin from Eastern Cottonmouth (*Agkistrodon piscivorus*) [35] and triflin from Habu snake (*Trimeresurus flavoviridis*) [36], with each of these proteins named after the snake specie in which it was discovered. It was reported that triflin reduced high potassium-induced smooth muscle contraction and blocked the L-type calcium channels [34], another reason that venoms were toxic.

3. Conclusions

In this study, we provided both electrophysiological and proteomic investigations to the toxicities and protein contents of *Protobothrops mangshensis* venom. Among the total of 143 proteins identified by MS/MS analysis of the tryptic peptides of the protein spots from 2-DE separation, ~59.4% proteins were various enzymes and ~24.5% were ion channel proteins. Many enzymes were found and identified to be with multiple isoforms observed in 2-DE. These results demonstrated a great variability of the snake venom proteins and enzymes, and correlated with high biological activities of the snake venom fluid, contributing to its extremely complex and highly toxic properties. Although many of the bioactivities of the identified proteins were based on similarity comparison of protein sequences and phylogenetic analysis, these results together with the electrophysiological study provided the first functional proteomics of *Protobothrops mangshensis* venom.

4. Supplementary material

Supplementary data and information is available at:

<http://www.jiomics.com/index.php/jio/rt/suppFiles/114/0>

Supplementary material includes Table 1 showing the list of the snake venom proteins identified by MS analysis.

Acknowledgements

This work was financially supported by the National Science Foundation of China (No. 21172067, No. 30971570, No. 31071091 and No. 31171196, the Science and Technology Department of Hunan Province (No. 2012SK3042), the Key Project of Education Department of Hunan, China (No. 09A035), and the Key Laboratory of Protein Chemistry and Developmental Biology of the Ministry of Education, Hunan Normal University. The specific thanks were conveyed to Mr. YuanHui Chen at the Department of Snake Research, Mangshan Forest Park, for his supplying *Photobothrops mangshanensis* crude venom.

References

1. Fox, J.W., Serrano, S.M.T., Snake venoms and coagulopathy. *Toxicon* 2005, 45, 951-967.
2. Cushman, D.W., Ondetti M.A. Design of angiotensin converting enzyme inhibitors. *Nature Medicine* 1999, 5, 1110-1113.
3. O'Shea, J.C., and Tchong, J.E., Eptifibatid: a potent inhibitor of the platelet receptor integrin glycoprotein Iib/IIIa. *Expert Opinion on Pharmacotherapy* 2002, 3, 1199-1210.
4. Burger, K.M., and Tuhim S., Antithrombotic trials in acute ischaemic stroke: a selective review. *Expert Opinion on Emerging Drugs* 2004, 9, 303-312.
5. Toombs, C. F., Alfineprase: pharmacology of a novel fibrinolytic metalloproteinase for thrombolysis. *Haemostasis* 2001, 31, 141-147.
6. Escoubas, P., Mass spectrometry in toxinology: a 21st-century technology for the study of biopolymers from venoms. *Toxicon* 2006, 47, 609-613.
7. Escoubas, P., Corzo, G., Whiteley, B.J., C  lerier, M.L., Nakajima, T., Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and high-performance liquid chromatography study of quantitative and qualitative variation in tarantula spider venoms. *Rapid Communications in Mass Spectrometry* 2002, 16, 403-413.
8. Favreau, P., Menin, L., Michalet, S., Perret, F., Cheneval, O., St  cklin, M., Bulet, P., St  cklin, R., Mass spectrometry strategies for venom mapping and peptide sequencing from crude venoms: case applications with single arthropod specimen. *Toxicon* 2006, 47, 676-687.
9. Legros, C., Celerier, M.L., Henry, M., Guette, C., Nanospray analysis of the venom of the tarantula *Theraphosa leblondi*: a powerful method for direct venom mass fingerprinting and toxin sequencing. *Rapid Communications Mass Spectrometry*. 2004, 18, 1024-1032.
10. Guette, C., Legros, C., Tournois, G., Goyffon, M., C  lerier, M.L., Peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of the *Lasiodora parahybana* tarantula venom gland. *Toxicon* 2006, 47, 640-649.
11. Richardson, M., Pimenta, A.M., Bemquerer, M.P., Santoro, M.M., Beirao, P.S., Lima, M.E., Figueiredo, S.G., Bloch, C.Jr., Vasconcelos, E.A., Campos, F.A., Gomes, P.C., Cordeiro, M.N., Comparison of the partial proteomes of the venoms of Brazilian spiders of the genus *Phoneutria*. *Comparative Biochemistry Physiology Toxicology & Pharmacology* 2006, 142, 173-187.
12. Zhao, E.M., Chen, Y.H., Description of a new species of the genus *Trimeresurus*. *Sichuan Journal of Zoology* 1990, 9, 11-12 (in Chinese).
13. David, P., Tong, H., Translations of recent descriptions of Chinese pitvipers of the *Trimeresurus* complex (Serpentes Viperidae), with a key to the complex in China and adjacent areas. *Smithsonian Herpetological Information Service* 1997, 112, 1-29.
14. Wuster, W., Peppin, L., Pook, C.E., Walker, D.E., A nesting of vipers: Phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes). *Molecular Phylogenetics Evolution* 2008, 49, 445-459.
15. Gumprecht, A., Tillack, F., Proposal for a replacement name of the snake genus *Ermia*. *Russian Journal of Herpetology* 2004, 11, 73-76.
16. Castoe, T.A., Parkinson, C.I., Bayesian mixed model and the phylogeny of pitvipers (Squamata: Serpentes). *Molecular Phylogenetics Evolution* 2006, 39, 91-110.
17. Mebs, D., Kuch, U., Coronas, F.I., Batista, C.V., Gumprecht, A., Possani, L.D., Biochemical and biological activities of the venom of the Chinese pitviper *Zhaoermia mangshanensis*, with the complete amino acid sequence and phylogenetic analysis of a novel Arg49 phospholipase A2 myotoxin. *Toxicon* 2006, 47, 797-811.
18. Yi, J., Craft, D., Gelfand, C.A., Minimizing preanalytical and analytical variability of plasma peptidomics using MALDI-TOF MS. *Methods of Mol. Biol.* 2011, 728, 137-49.
19. Yi, J., Changki K., Gelfand, C.A., Inhibition of Intrinsic Proteolytic Activities Moderates Preanalytic Variability and Stability of Human Plasma, *J. of Proteome Res.*, 2007, 6, 1768-81.
20. Huang, L.Y., Li, B.X., Luo, C., Xie, J.Y., Chen, P., Liang, S.P., Proteome comparative analysis of gynogenetic haploid and diploid embryos of goldfish (*Carassius auratus*). *Proteomics* 2004, 4, 235-243.
21. Hellmann, U., Wemstedt, C., Gonez, J., Heldin, C.H., Improvement of an "In-Gel" digestion procedure for the micro-preparation of internal protein fragments for amino acid sequencing. *Analytical Biochemistry* 1995, 224, 451-455.
22. Cao, R., Li, X., Liu, Z., Liang, S.P. Integration of a two-phase partition method into proteomics research on rat liver plasma membrane proteins. *Journal of Proteome Research* 2006, 5, 634-642.
23. Gobom, J., Schuerenberg, M., Mueller, M., Theiss, D., Leh-rach, H., Nordhoff, E., Alpha-cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics. *Analytical Chemistry*. 2001, 73, 434-438.
24. Amiconi, G., Amoresano, A., Boumis, G., Brancaccio, A., De Cristofaro, R., De Pascalis, A., Di Girolamo, S., Maras, B., Scaloni A., A novel venombin B from *Agkistrodon contortrix contortrix*: evidence for recognition properties in the surface around the primary specificity pocket different from thrombin. *Biochemistry* 2000, 39, 10294-10308.
25. Matsui, T., Sakurai, Y., Fujimura, Y., Hayashi, I., Oh-Ishi, S., Suzuki, M., Hamako, J., Yamamoto, Y., Yamazaki, J., Kinoshita, M., Titani K., Purification and amino acid sequence of halystase from snake venom of *Agkistrodon halys blomhoffii*, a serine protease that cleaves specifically fibrinogen and kininogen. *European Journal of Biochemistry* 1998, 252, 569-575.
26. Atoda, H., Yoshida, N., Ishikawa, M., Morita, T., Binding

- properties of the coagulation factor IX/factor X-binding protein isolated from the venom of *Trimeresurus flavoviridis*. *European Journal of Biochemistry* 1994, 224, 703-708.
27. Hirabayashi, J., Kusunoki, T., Kasai, K., Complete primary structure of a galactose-specific lectin from the venom of the rattlesnake *Crotalus atrox*. Homologies with Ca²⁺-dependent-type lectins. *Journal of Biology Chemistry* 1991, 266, 2320-2326.
 28. Chen, R.Q., Jin, Y., Wu, J.B., Zhou, X.D., Lu, Q.M., Wang, W.Y., Xiong, Y.L., A new protein structure of P-II class snake venom metalloproteinases: it comprises metalloproteinase and disintegrin domains. *Biochemical Biophysical Research Communications* 2003, 310, 182-7.
 29. Du, X.Y., Clemetson, K.J., Snake venom L-amino acid oxidase isoenzymes. *Toxicon* 2002, 40, 659-65.
 30. Huang, M.Z., Gopalakrishnakone, P., Chung, M.C.M., Kini, R.M., Complete amino acid sequence of an acidic, cardiotoxic phospholipase A2 from the venom of *Ophiophagus Hannah* (king cobra): a novel cobra venom enzyme with "pancreatic loop". *Archives Biochemistry Biophysics* 1997, 338, 150-6.
 31. Tsai, I.H., Chen, Y.H., Wang, Y.M., Comparative proteomics and subtyping of venom phospholipases A₂ and disintegrins of *Protobothrops pitvipers*. *Biochimica Biophysica Acta* 2004, 1702, 111-119.
 32. Mebs, D., Kuch, U., Coronas, F.I., Batista, C.V., Gumprecht, A., Possani, L.D., Biochemical and biological activities of the venom of the chinese pitviper *Zhafermia mangshanensis*, with the complete amino acid sequence and phylogenetic analysis of a novel Arg49 phospholipase A2 mytotoxin. *Toxicon* 2006, 47, 797-811.
 33. Yamazaki, Y., Morita, T., Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 2004, 44, 227-231.
 34. Yamazaki, Y., Koike, H., Sugiyama, Y., Motoyoshi, K., Wada, T., Hishinuma, S., Mita, M., Morita, T., Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. *European Journal Biochemistry* 2002, 269, 2708-2715.
 35. Yamazaki, Y., Hyodo, F., Morita, T., Wide Distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins. *Archives Biochemistry Biophysics* 2003, 412, 133-141.
 36. Shikamoto, Y., Suto, K., Yamazaki, Y., Morita, T., Mizuno, H., Crystal structure of a CRISP family calcium-channel blocker derived from snake venom. *Journal Molecular Biology* 2005, 350, 735-743.