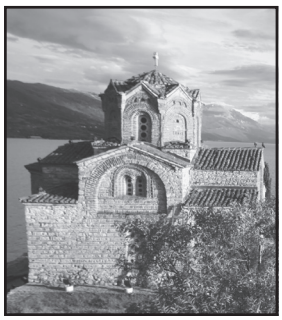
BJMG
VOL. 15, 2012
Supplement
ISSN 1311-0160

Balkan Journal of Medical Genetics

International Journal of Medical Genetics
Supplement

MACPROGEN Final Conference
Ohrid, 2012



BJMG is an open access, peer-reviewed journal. All articles in the journal are immediately and permanently available online for all to read and use free of charge.
BJMG is published twice a year.
Indexed or abstracted in: EMBASE (the Excerpta Medica database), Elsevier Biobase (Current Awareness in Biological Sciences), Elsevier GeoAbstracts, Scopus, Chemical Abstracts (CA), Science Citation Index Expended, Journal Citation Reports/Science Edition, VERSITA Central European Science Publishers and Copernicus.
BJMG is covered by Thomson Reuters services.

EDITORS

Plaseska-Karanfilska D (Skopje), Toncheva DI (Sofia)

PAST EDITOR

Efremov GD (Skopje)

EDITORIAL BOARD

Apak M (Istanbul)
Avent N (Plymounh)
Baranov V (St Petersburg)
Barisic I (Zagreb)
Basak N (Istanbul)
Czeizel A (Budapest)
Cooper D (Cardiff)
Dimovski AJ (Skopje)
Dörk T (Hannover)
Efremov DG (Rome)
Frins JP (Leuven)
Geraedts J (Maastricht)
Goossens M (Paris)
Kanavakis E (Athens)
Keohavong P (Pittsburg)
Kocova M (Skopje)
Kremenski I (Sofia)
Kucinskas V (Vilnius)
Khusnutdinova E (Ufa)
Labie D (Paris)
Liehr T (Jena)
Lungeanu A (Bucharest)
Nakamara Y (Tokyo)
Pavelic K (Zagreb)
Peterlin B (Ljubljana)
Pignatti PF (Verona)
Stevanovic M (Belgrade)
Todorova A (Sofia)
Tomlinson I (London)
Vella F (Saskatoon)
Wajcman H (Paris)
Weatherall D (Oxford)
Wong S (Hong Kong)
Yapiyakis C (Athens)
Zoraqi G (Tirana)

EDITORIAL ASSISTANT

Sukarova Stefanovska E (Skopje)
Carver M.F.H (Evans, GA)
Trencevska B (Skopje)

COVER DESIGN

Bitrakova-Grozdanova V (Skopje)

EDITORIAL OFFICES

Macedonian Academy of Sciences and Arts,
Av. Krste Misirkov 2, POB 428,
1000 Skopje, Republic of Macedonia.
Phone: + 389 2 3235 411, fax: + 389 2 3115 434, E-mail: manu.icgib@gmail.com
Web page: www.bjmg.edu.mk

PUBLISHER

BJMG is published by Macedonian Academy of Sciences and Arts
Printed by Jugoreklam – Skopje, Republic of Macedonia

Front cover page: Saint John the Theologian, Kaneo, Macedonian Ortodox Church, 13th Century, Ohrid, Republic of Macedonia

Back cover page: Participants of MACPROGEN Final Conference at St Naum's Monastery, Ohrid, Republic of Macedonia

PROCEEDINGS

of the MACPROGEN Final Conference

held at Ohrid, Republic of Macedonia
March 29 – April 1, 2012

29 March – 1 April, 2012
Ohrid, Republic of Macedonia

Conference Programme

Thursday, 29 March 2012

13:00—17:00 ***Registration***

17:00—17:30 Welcome Address
Momir Polenakovic, RCGEB, Skopje, Republic of Macedonia
Dijana Plaseska-Karanfilska, RCGEB, Skopje, Republic of Macedonia
MACPROGEN Overview

SESSION 1 ***Chairpersons: Momir Polenakovic and Max Chaffanet***

17:30—18:00 **Massimo Delledonne,** University of Verona, Verona, Italy
Microarray or RNA Sequencing: what to use if you are "just" a biologist

18:00—18:30 **Martina Marchetti-Deschmann,** University of Technology, Vienna Austria
Renopathological microstructure visualization from formalin fixe kidney tissue by MALDI TOF imaging mass spectrometry

Friday, 30 March 2012

SESSION 2 ***Chairpersons: Borut Peterlin and Aleksandar Dimovski***

09:00—09:30 **Neil Avent,** Plymouth University, Plymouth, Devon, UK
Non invasive prenatal diagnosis of aneuploidy — next-generation sequencing or fetal DNA enrichment?

09:30—10:00 **Rafael Oliva,** University of Barcelona, Barcelona, Spain
Proteomics of the spermatozoon

10:00—10:30 **Dijana Plaseska-Karanfilska,** RCGEB, Skopje, Republic of Macedonia
Genetic causes of male infertility

10:30—10:45 **Katarina Davalieva,** RCGEB, Skopje, Republic of Macedonia
Differential expression profile of seminal plasma protein in men with spermatogenic impairment

10:45—11:00 **Svetlana Madjunkova,** RCGEB, Skopje, Republic of Macedonia
Prenatal diagnosis — rapid and non invasive approaches

SESSION 3 ***Chairpersons: Rafael Oliva and Natalia Bogdanova***

11:30—12:00 **Borut Peterlin,** University of Ljubljana, Ljubljana, Slovenia
Integratomic approach to gene candidate identification

12:00—12:30 **Cristina Patuzzo,** University of Verona, Verona, Italy
A preliminary microRNA analysis of thoracic aortic aneurysms

12:30—13:00 **Aleksandar Dimovski,** University of Skopje, Skopje, Republic of Macedonia
Personalized medicine: scientific, ethical and legal aspects

13:00—13:15 **Emilija Sukarova Stefanovska,** RCGEB, Republic of Macedonia
Genetics of nonsyndromic deafness in the Republic of Macedonia

13:15—13:30 **Biljana Atanasovska,** RCGEB, Skopje, Republic of Macedonia
Molecular diagnostics of Mediterranean β-thalassemia

13:30—13:45 **Sanja Kiprijanovska,** RCGEB, Skopje, Republic of Macedonia
Study of hepatitis C virus infections in the Republic of Macedonia

Saturday, 31 March 2012

SESSION 4 ***Chairpersons: Cristina Patuzzo and Andreja Arsovski***

15:30—16:00 **Max Chaffanet,** Institute Paoli Calmettes, Marseilles, France
Integrated genomic alteration in breast cancer

16:00—16:30 **Natalia Bogdanova,** Hannover Medical School, Hannover, Germany
Molecular genetics of breast and ovarian cancer:
new developments and therapeutic implications

16:30—16:45 **Ivana Maleva,** RCGEB, Skopje, Republic of Macedonia
Genetic basics of breast cancer in the Republic of Macedonia

16:45—17:00 **Katarina Popovska-Jankovic,** RCGEB, Skopje, Republic of Macedonia
MicroRNAs in breast cancer — our initial results

SESSION 5 ***Chairpersons: Martina Marchetti-Deschmann and Neil Avent***

17:30—18:00 **Bruno Domon,** Luxembourg Clinical Proteomics Center, CRP-Santé, Luxembourg
Recent developments in mass spectrometry-based proteomics approaches

18:00—18:30 **Niels Tommerup,** Wilhelm Johannsen Centre, University of Copenhagen, Copenhagen, Denmark
Chromosomal breakpoints target all parts of the human genome

18:30 **Closing Remarks**

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

WELCOME ADDRESS

Respected Academicians,
Distinguished Guest Speakers,
Colleagues,
Ladies and Gentlemen

It is my great pleasure to welcome you to the Final MACPROGEN Conference. To begin, I would like to take this opportunity to thank the European Commission for their support. I would also like to express our gratitude to the guest speakers, who represent our partners in the MACPROGEN project. We are grateful for your collaboration; your expertise was a valuable contribution to the successful realization of the MACPROGEN project. I would also like to thank our collaborators from different Macedonian institutions for their participation in this Conference and for their continuous support and collaboration. We are honored to have you all here with us.

Unfortunately, during the work on the MACPROGEN Project we lost our professor and mentor, Georgi D. Efremov, who was a founder and a Director of the Research Centre for Genetic Engineering and Biotechnology (RCGEB), since its establishment until last May when he passed away. Professor Efremov was a world renowned expert in the field of hemoglobin (Hb) research and a very prominent expert in the field of biomedicine and biomo-lecular sciences in the Republic of Macedonia and in the surrounding regions. Immediately after his death, the Presidency of the Macedonian Academy of Sciences and Arts renamed the Center in his honor as the Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov." I would just like to emphasize that the European Commission Research Potential grants were very competitive; our Institution was successful primarily because Academician Efremov had devoted his whole professional career and life to build such a prestigious Center with a high international reputation.

I hope that in the next 3 days in this beautiful city of Ohrid we will have a successful conference full of interesting and productive discussions and exchange of ideas, but also some free time to enjoy the sightseeing of the ancient city of Ohrid.

At the end of the MACPROGEN Project, I would like to express my belief that this is just the end of one project, but at the same time, a start of many more.

Dijana Plaseska-Karanfilska
MACPROGEN Coordinator

List of Participants

(in alphabetical order)

Arsovski Andrea	R. Macedonia
Atanasovska Biljana	R. Macedonia
Avent Neil	United Kingdom
Bogdanova Natalia	Germany
Bozinovski Gjorgji	R. Macedonia
Chaffanet Max	France
Chakalova Ljubomira	R. Macedonia
Davalieva Katarina	R. Macedonia
Delledonne Massimo	Italy
Dimovski Aleksandar	R. Macedonia
Domon Bruno	Luxembourg
Duvlis Sotirija	R. Macedonia
Janeva Stanka	R. Macedonia
Karanfilski Oliver	R. Macedonia
Kiprijanovska Sanja	R. Macedonia
Koceva Svetlana	R. Macedonia
Kubelka Katerina	R. Macedonia
Madjunkova Svetlana	R. Macedonia
Maleva Ivana	R. Macedonia
Marchetti-Deschmann Martina	Austria
Matevska Nadica	R. Macedonia
Mircevska Marija	R. Macedonia
Moneva Zvezdana	R. Macedonia
Noveski Predrag	R. Macedonia
Oliva Rafael	Spain
Patuzzo Cristina	Italy
Petanovski Zoranco	R. Macedonia
Peterlin Borut	Slovenia
Plaseska-Karanfilska Dijana	R. Macedonia
Polenakovic Momir	R. Macedonia
Popov Zivko	R. Macedonia
Popovska Jankovic Katerina	R. Macedonia
Simjanovska Liljana	R. Macedonia
Stefanovski Kliment	R. Macedonia
Sukarova Stefanovska Emilija	R. Macedonia
Tommerup Niels	Denmark
Trencevska Blagica	R. Macedonia
Trpevski Mirko	R. Macedonia

CONTENTS

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

NATIONAL REFERENCE CENTRE FOR GENOMICS AND PROTEOMICS — MACPROGEN

Plaseska-Karanfilska D*

***Coordinator of the MACPROGEN Project:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

INTRODUCTION

The "National Reference Centre for Genomics and Proteomics" (MACPROGEN) was financed by the European Commission within the FP7 Capacities Work Program for a period of 3 years, starting April 1 2009. The MACPROGEN project aimed at upgrading and improving the capacity of the Research Centre for Genetic Engineering and Biotechnology (RCGEB) "Georgi D. Efremov," Macedonian Academy of Sciences and Arts (MASA), Skopje, Republic of Macedonia, for research and education in the fields of genomics and proteomics. The main objectives of the project were to establish a technological platform for high throughput genomics and proteomics research, networking with European Union (EU) research institutions in order to foster collaborative activities, disseminating knowledge and expertise and ultimately building an interactive and competitive research environment. Altogether, the realization of MACPROGEN objectives was contributing to the strengthening of Macedonian research, technology and development capacities and enhancing career opportunities for Macedonian scientists.

The RCGEB is one of the few institutions from the Republic of Macedonia that met the criteria for a center of excellence due to its esteemed scientific and educational record in life sciences. The RCGEB was founded in 1987 as a research unit of MASA with the main goal of advancing scientific knowledge in the field of protein chemistry, molecular biology, genetic engineering and biotechnology through research, practical training of scientists and postgraduate studies. Under the guidance of the late Academician Georgi D. Efremov, the founder and Director of RCGEB until his death in May 2011, the RCGEB became a hub for research in the field of biomolecular sciences in the Republic of Macedonia. It was one of the first institutions in the region that applied these new technologies in molecular diagnostics of human diseases and became an international center for training in basic and advanced methods in these sciences. Immediately after his death, the Presidency of MASA renamed the center in his honor, as the Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov." In the past 25 years, the RCGEB scientists have published over 150 papers in international and national journals, as well as several chapters in books and monographs. The RCGEB collaborates with numerous institutions from the Republic of Macedonia, medical institutions from neighboring countries and many academic institutions from different countries worldwide.

The main research interest at the RCGEB during the past 25 years has been molecular characterization of the most common monogenic diseases, with a special emphasis on hemoglobinopathies, as well as some aspects of the molecular epidemiology of

infectious diseases, genetics of the most common malignancies and DNA markers for human identification. The upgrading of the RCGEB infrastructure by MACPROGEN funds has enabled widening of the research objectives toward larger scale investigations of monogenic diseases and also shifting our research interest toward comprehensive studies of some common complex diseases such as cancers, infertility, mental retardation and deafness by high throughput genomic and proteomic technologies.

Keywords: Genomics, Proteomics, Republic of Macedonia

MACPROGEN PROJECT ACTIVITIES

The activities within the MACPROGEN project were organized into *five work packages*: Management and Coordination (WP1), Technological Platform and Employment (WP2), Networking and Training (WP3), Workshops and Symposia (WP4) and Promotion and Dissemination (WP5).

WP1: Management and Coordination covered administrative, financial and technical management of the project activities. The activities within this project were coordinated by the Steering Committee which was composed of the *Coordinator* of the project and the *Work Package Leaders*. The Steering Committee worked closely with the administration and financial departments of MASA, as well as the RCGEB Scientific Committee. The Steering Committee met regularly each month to enable a smooth operation of all foreseen activities and to fulfill the objectives of the project.

WP2: Technological Platform and Employment was headed by the late Academician Georgi D. Efremov (until May 2011) and by Professor Dr. Dijana Plaseska-Karanfilska (since May 2011). The three main objectives of WP2: ***1)*** purchasing of equipment; ***2)*** preparation of working protocols and operational use of equipment; and ***3)*** hiring of scientists, were all successfully completed. The following equipment was purchased: DNA microarray system, 2-D DIGE (two-dimensional differential in gel electrophoresis) system, Genetic Analyzer, real-time polymerase chain reaction (ReTi-PCR) system, Bioanalyzer, spectrophotometers, microcentrifuges and liquid scintillation counter. The equipment has already been integrated into RCGEB laboratory practices and numerous working protocols have been developed using the new equipment. The upgrading of the RCGEB infrastructure was also supported by the Government of the Republic of Macedonia by purchasing additional equipment: 2-D nano HPLC (high performance liquid chromatography), Accuspot and a matrix-assisted laser desorption/ionization time of flight-time of flight (MALDI TOF-TOF) mass spectrometer.

Nine researchers, paid by MACPROGEN funds, have been employed at RCGEB during the work on this project: Lybomira Chakalova, Ph.D., Svetlana Madzunkova, M.D., M.Sc., Predrag Noveski, M.Sc., Sanja Kiprijanovska, Ivana Maleva, M.Sc., Biljana Atanasovska, Zvezdana Moneva, M.Sc., Slavica Pecioska M.Sc and Ognen Spiroski M.Sc. These researchers have become an integral part of the research potential of RCGEB. The Presidency of MASA in coordination with RCGEB is making efforts to obtain permanent positions paid for by the Government of the Republic of Macedonia for the researchers who were working on the MACPROGEN project. Until this is realized, RCGEB has ensured that funds are available for their salaries for the period after completion of the MACPROGEN project.

WP3: Networking and Training was headed by Professor Dr. Dijana Plaseska-Karanfilska. The main objectives of this work package were to: ***1)*** foster networking and establish close collaboration with leading EU institutions; ***2)*** transfer of the high throughput genomic and proteomic technologies and scientific knowledge; and ***3)*** foster preparation of collaborative projects.

Six leading EU institutions were our partners since the beginning of the project: ***1)*** University of Copenhagen, Wilhelm Johannsen Centre for Functional Genome Research, Copenhagen, Denmark; ***2)*** Hannover Medical School, Gynaecology Research Unit, Hannover, Germany; ***3)*** University of Barcelona, Faculty of Medicine, Department of Physiological Sciences, Human Genetics Laboratory, Barcelona, Spain; ***4)*** University of Verona, Department of Mother and Child, Biology and Genetics, Section of Biology and Genetics, Verona, Italy; ***5)*** University of Copenhagen, Division of Genetics and Bioinformatics, IBHV, Copenhagen, Denmark; and ***6)*** Institute Paoli Calmettes, Molecular Oncology Department, Oncogenomic Group, Marseille, France.

However, during the work on this project we have established close collaboration with other institutions, namely: ***1)*** Mondor Institute of Biomedical Research, Henri Mondor Hospital, Créteil, France; ***2)*** Institut Cochin, Inserm, University Paris Descartes, Paris, France; ***3)*** K.U. Leuven, Labaratory for Cytogenetics and Genome Research, Department of Human Genetics, Biomedical Sciences Group, Leuven, Belgium; ***4)*** Vienna University of Technology, Institute of Chemical Technologies and Analytics, Vienna, Austria; ***5)*** Clinical Proteomics Center, Luxemburg; ***6)*** Columbia University, Department of Pathology and Cell Biology, New York, NY, USA; ***7)*** Mabritec AG, Riehen, Switzerland; ***8)*** University Medical Centre Ljubljana, Department of Obstetrics and Gynecology, Division of Medical Genetics, Ljubljana, Slovenia; ***9)*** Bulgarian Academy of Sciences, Institute of Experimental Morphology and Anthropology with Museum Sofia, Bulgaria; ***10)*** University of Glasgow, College of Medical, Veterinary and Life Sciences, BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, Glasgow, Scotland, UK; and ***11)*** Academy of Athens, Biomedical Research Foundation, Biotechnology Division, Athens, Greece.

Knowledge transfer was achieved through the training of MACPROGEN researchers in the new technologies. Several visits of researchers from RCGEB to EU laboratories were organized. These proposed to ensure that RCGEB researchers are familiarized with the new technologies [microarray, 2-D DIGE, mass spectrometry (MS)] through practical experience. Several expert visits were also organized. Some of them were primarily aimed to establish closer professional contacts and facilitate collaboration. In other cases, the main objective was to present and promote MACPROGEN along with the RCGEB. Since the starting date of the project, researchers on the MACPROGEN team have applied for a total of 17 research projects, of which, seven were successful and four are awaiting decisions. Work on funded projects has commenced in accordance with the respective grant agreements.

WP4: Workshops and Symposia was headed by Dr. Katarina Davalieva. The objective of WP4 was the organization of three Genomics and Proteomics Workshops and a Final Conference. The first workshop was organized in November 2010, the second in June 2011 and the third in March 2012. The theoretical part of the workshops included lectures given by scientists from the partnering institutions; the lectures were attended by the members of the Macedonian Biochemical Society and Macedonian Society of Human Genetics. The practical part of all three workshops was mainly attended by RCGEB scientists and collaborators, but on the first workshop 24 scientists from different Macedonian research institutions also participated. The practical part of all three workshops included demonstration of protocols using the major new equipment.

The Final Conference took place at Ohrid, Republic of Macedonia, at the end of March 2012. Ten scientists from 10 different EU institutions participated, in addition to all RCGEB employees and several collaborators from the Republic of Macedonia.

WP5: Promotion and Dissemination was headed by Dr. Emilija Sukarova-Stefanovska. The objectives of this work package were to: ***1)*** promote the activities and potential of the National Reference Centre for Genomics and Proteomics; ***2)*** increase and strengthen the collaboration with institutions from the country, the wider region and the EU; and ***3)*** increase the participation of scientists from the Republic of Macedonia in FP7. These objectives have been accomplished as planned. The website of the National Reference Centre for Genomics and Proteomics (www.manu.edu.mk/macprogen) was launched during the second month and has been regularly updated to reflect the progress of project activities. Leaflets containing relevant information on MACPROGEN have been published three times and distributed during events, such as meetings, lectures, *etc.* Meetings with clinicians, health professionals and research scientists from several institutions have been periodically organized to discuss the Centre's policies for achieving optimal translation of its potential. During the work on the MACPROGEN project, RCGEB scientists have participated in several scientific meetings and published 36 conference abstracts. The MACPROGEN researchers have also published 25 papers in international and national journals.

Project Impact. The introduction of novel genomic and proteomic technologies at RCGEB

"Georgi D. Efremov" has already resulted in strengthening of Macedonian and regional research capacities. Experienced scientists, some of them trained in EU countries, have joined the RCGEB team. Contemporary research projects in collaboration with leading EU institutions have been initiated. It is very much hoped that the modernization of research along with intensified collaboration will further stimulate promising young scientists to establish their research in the Republic of Macedonia. It is expected that the improved infrastructure and research environment will have long-lasting effects, such as enhanced levels of science communication and high-impact publications.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

RENOPATHOLOGICAL MICROSTRUCTURE VISUALIZATION FROM FORMALIN FIXED KIDNEY TISSUE BY MATRIX-ASSISTED LASER/DESORPTION IONIZATION-TIME-OF-FLIGHT MASS SPECTROMETRY IMAGING

Fröhlich S[1], Putz B[1], Schachner H[2], Kerjaschki D[2],Allmaier G[1], Marchetti-Deschmann M[1,*]

***Corresponding Author:** Dr. Martina Marchetti-Deschmann, Vienna University of Technology, Institute of Chemical Technologies and Analytics, Getreidemarkt 9/164-IAC, 1060 Vienna, Austria; Tel.: +43-1-58801-15152; Mobile: +43+664-605887663; Fax: +43-1-58801-915162; E-mail: martina.marchetti-deschmann@tuwien.ac.at

ABSTRACT

Understanding early stage renal malfunctions with regard to the glomerular filtration processes is essential for nephropathological prescreening strategies and intervention at an early stage. Mass spectrometry imaging (MSI) in combination with histopathology can provide an universal analytical approach. Proteomic and lipidomic aspects of glomerular biocompositions were applied for microstructural differentiation in healthy rat kidney samples. Usability of commonly used tissue embedding media and the compatibility of histological staining and fixation methods were of interest. It was demonstrated that ultra-thin tissue samples (500 nm, 1 and 10 µm) can be used for lipid and peptide-based differentiation at the glomerular resolution level in formalin-fixed tissue samples in combination with preceding histological staining for correlating optical and molecular mass images.

Keywords: Glomerular microstructures, Lipids, Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS) imaging, Nephropathy, Proteins

[1] Vienna University of Technology, Institute of Chemical Technologies and Analytics, Vienna, Austria

[2] Medical University of Vienna, Department of Clinical Pathology, Vienna, Austria

INTRODUCTION

A combination of histopathology and mass spectrometry imaging (MSI) offers comprehensive information regarding structure, molecular composition and pathological information in tissue samples. Mass spectrometry imaging is a rapidly developing technique using spatially resolved mass spectrometry (MS) techniques to simultaneously trace distributions of hundreds of biomolecules directly from tissue samples using essentially the same technology. Peptides, proteins, pharmaceuticals and metabolites can also be analyzed but without a label and without prior knowledge. With MSI, molecular peak information is correlated to the underlying tissue architecture and a virtual image is rebuilt with respect to the intensity of each molecular species to understand the distribution of differential signals.

Renal dysfunction has a high demand for early stage identification to initiate healing processes. Recent investigations proved nephropathy is associated with insulin resistance leading to malfunctions of the glomerular filtration [1], manifested in potential glomerular basement membrane modifications [2]. Lipid phosphatase levels, promoting podocyte apoptosis leading to diabetic nephropathy, are upregulated before histological changes are observed [1]; moreover, the membrane contains protein complexes associated with cholesterol binding which alters the lipid environment [2]. Lipid accumulations

have been reported for diabetic kidney diseases [3] and oxidized phosphatidylcholine species seem to be associated with renal dysfunction [4]. Recently tubuli-related phosphatidylcholine classes identified by MSI were correlated to immunoglobulin A nephropathy [5].

Thus, for visualizing renopathological microstructures, MSI is a promising tool, allowing further discoveries of yet unknown disease-related analytes. Special challenges regarding ion suppression effects during the matrix-assisted laser desorption/ionization (MALDI) process were observed whilst obtaining lipidomic information using MSI. Consequently, special attention had to be paid to sample preparation methods, regarding washing procedures, matrix application and protein denaturation, making adaptation to the particular analytical question mandatory [6].

MATERIALS AND METHODS

Chemicals were of analytical grade (Sigma-Aldrich, St. Louis, MO, USA). Water (ddH_2O) was purified in a Simplicity system (Millipore, Billerica, MA, USA). Frozen rat kidney tissue (formalin-fixed, unfixed) embedded in paraffin (FFPE), optimal cutting temperature (OCT) compound (Sakura Finetek, Tokyo, Japan) or sucrose was sliced to 10, 5 and 0.5 µm using a cryo-microtome (Leica, Wetzlar, Germany) and mounted on indium-tin oxide coated glass slides (Sigma-Aldrich). The OCT compound and sucrose were removed by washing the samples five times (4°C) for 45 seconds with 200 µL ddH_2O/cm^2. A xylene bath followed by a descending ethanol gradient (100.0, 96.0, 70.0, 50.0, 0.0% in ddH_2O) removed the paraffin. All samples were washed three times with 200 µL 70.0% ethanol/cm^2 for 45 seconds (4°C) and vacuum dried for 15 min. before the MALDI matrix application. Samples were hematoxylin/eosin (H/E) stained before MSI treatment.

The MALDI matrices [α-cyano-4-hydroxycinnamic acid (HCCA), sinapinic acid, 2,5-dihydroxybenzoic acid] were dissolved in solvents containing 50.0 or 70.0% acetonitrile or ethanol in 0.1% aqueous trifluoroacetic acid. Matrix deposition was performed using a ChIP-1000 (Shimadzu Biotech Kratos Analytical). For protein identification, tissue was trypsinized before matrix deposition by depositing enzyme directly on pre defined tissue spots [1 ng/µL in 50 mM ammonium bicarbonate, 0.1% Rapigest (Waters, Manchester, Greater Manchester, UK)]. Samples were incubated in a humidified atmosphere overnight at 37°C before desiccation in vacuum and heat treatment (85°C).

Mass spectrometry imaging experiments were performed using a MALDI-TOF (time of flight) instrument (AXIMA TOF^2; Shimadzu Biotech Kratos Analytical, Manchester, Greater Manchester, UK; 337nm nitrogen laser, 20Hz) and a MALDI-QqTOF instrument (Synapt HDMS, Waters; 355nm Nd-YAG laser, 200Hz). The optical diameter of 60 µm was reduced to an operational resolution of 35 µm by over sampling.

Peptide and lipid identification was based on post source decay (PSD) and collision induced dissociation (CID) fragmentation in combination with database search [proteins: SwissProt (http://www.uniprot.org/help/ uniprotkb); lipids: Lipid Maps (http://www.lipidmaps.org/ data/structure/)]. Selected ion images were generated using BioMAP (Novartis, Basel, Switzerland).

RESULTS AND DISCUSSION

Histology and Mass Spectrometry Imaging. Integrating MSI in histopathology necessitates the combination of different requirements from pathology and MS. The primary demands for histology are the preservation of high spatial resolution for tissue structures, long lasting analyte preservation and low sample consumption. For MS, spatial resolution can usually be disregarded and analytes have to be mobilized instead of preserved. For MSI, additionally, histological information has to be retained and tissue disintegration has to be limited to a minimum. We present an approach arranging histological staining prior to MSI using ultra-thin tissue sections showing good molecular results comparable to microscopic images.

It was found that MALDI-TOF MSI analysis of H/E-stained tissue showed no limiting aspects concerning proteomic approaches. Protein digestion directly from the tissue and measuring peptides by MS could be achieved. Stain-related signals, however, were observed below m/z signals of 500, eventually leading to ion suppression effects for neutral lipids. Phospholipid classes were obtained unaffected from

ultra-thin sucrose samples. Histological images of OCT embedded samples were often difficult to correlate to MSI results because of structural changes after thawing. Sucrose embedding in combination with cutting thickness of 1 µm was sufficient for IMS analysis, even reducing signal background and enhancing signal quality. The selected ultra-thin samples considerably improved signal quality whilst preserving spatial resolution.

Lipid Differentiation. Lipid-based histological differentiation was performed using unfixed rat kidney samples embedded in sucrose and formalin-fixed tissue embedded in OCT Low embedding medium related signals and intensive species-related signals were obtained from both OCT and sucrose embedded samples. Several lipid classes could be localized, identified and characterized by fragmentation experiments. Unfixed tissue species revealed different distribution pattern phosphoatidylcholine species with varying alkyl chain lengths. The 1-tetradecanoyl-2-sn-glycero-3-phosphocholine, 1-hexadecanolyl-2-sn-glycero-3-phosphocholine and 1-octadecanoyl-2-sn-glycero-3-phosphocholine showed specifiable areas related to kidney structures, *i.e.,* cortex, medulla and pelvis (Figure 1). The 1-hexadecyl-2-acetyl-sn-glycero-3-phosphocholine and 1-hexadecanoyl-sn-glycero-3-phosphocholine exhibited different localization as well as 1-O-(1Z-tetradecenyl)-2-(9Z-octadecenoyl)-sn-glycerol and its oxidized co-species. Besides 16 different phospholipids, diacylglycerophosphoethanolamine, N-(tetracosanoyl)-sphing-4-enine-1-phospho-choline and two glomeruli specific metabolites, cholesterol and squalene, were also discovered. Histological correlation revealed the potential to identify pathologically modified membrane structures in glomerular associated diseases.

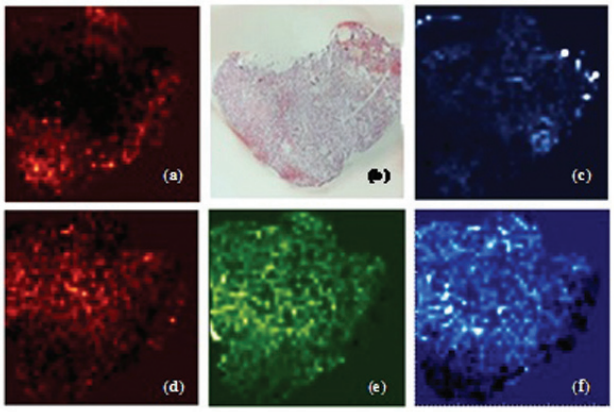


Figure 1. Tissue differentiation according to identified lipid species. Distribution of **(a)** diacylglycerol on **(b)** 5 µm OCT embedded unfixed kidney tissue; **(c)** oxidized diacylglycerol; **(d)** 1-tetradecanoyl-2-sn-glycero-3-phos-phocholine; **(e)** 1-hexadecanolyl-2-sn-glycero-3-phosphocholine; **(f)** 1-octadecanoyl-2-sn-glycero-3-phosphocholine.

Protein Identification. To detect most of the proteins available in an ultra-thin sample, several washing steps were required for complete removal of analyte species which are much easier to ionize than proteins, *e.g.,* lipids [7]. Our study showed that mass spectra quality for tissue analysis was enhanced by heat treatment before MSI experiments. Several tissue-related peptides could be identified by PSD and CID, revealing the possibility to distinguish pelvis and cortex, both tissue macrostructures. For all investigated embedding materials as well as fixation methods, it was possible to obtain sufficient signals up to m/z values of 10000 (Figure 2). The application of danaturation agents and heat treatment supports subsequent on-tissue enzymatic treatment for protein identification on fixed and embedded rat kidney. Highly cross-linked proteins (formalin) showed improved signal intensities for tryptic peptides after denaturation, which significantly supported protein identification based on peptide fragmentation. For protein identification in MSI approaches, peptide sequencing was indispensable because of protein co-localization and the non applicable approach of peptide mass fingerprinting. The presented protocol on ultra-thin tissue sections

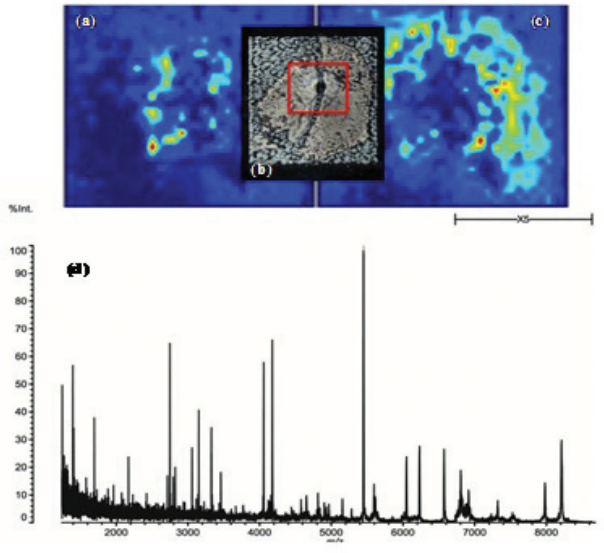


Figure 2. Protein based differentiation of **(a)** medulla and **(c)** cortex in; **(b)** FFPE rat kidney tissue (10 µm) after HCCA application; **(d)** MALDI-TOF-MS profile of the marked sample area (red).

allowed the identification of Tubulin and Claudin-4, besides seven other peptide fragments on differentiated, but so far unspecific, tissue regions.

Perspectives. Combining histopathology and MSI for investigating structural membrane modifications associated with glomerular related nephropathology looks promising. Structural modifications and molecular changes can be detected and investigated very early in progression.

REFERENCES

1. Hyvönen ME, Saurus P, Wasik A, *et al.* Lipid phosphotase SHIP2 downregulates insulin signalling in podocytes. Mol Cell Endocrinol. 2010; 328(1-2): 70-79.
2. Kibel G, Heilhecker A, Von Bruchhausen F. Lipids associated with bovine kidney glomerular basement membranes. Biochem J. 1976; 155(3): 535-541.
3. Wang Z, Jiang T, Li J, *et al.* Regulation of renal lipid metabolism, lipid accumulation, and glomerulosclerosis in FVBdb/db mice with type 2 diabetes. Diabetes. 2005; 54(8): 2328-2335.
4. Jimi S, Uesugi N, Saku K, *et al.* Possible induction of renal dysfunction in patients with lecithin:cholesterol acyltransferase deficiency by oxidized phosphatidylcholine in glomeruli. Arterioscler Thromb Vasc Biol. 1999; 19(3): 794-801.
5. Kaneko Y, Obata Y, Nishino T, *et al.* Imaging mass spectrometry analysis reveals an altered lipid distribution pattern in the tubular areas of hyper-IgA murine kidneys. Exp Mol Pathol. 2011; 91(2): 614-621.
6. Kaletaş BK, van der Wiel IM, Stauber J, *et al.* Sample preparation issues for tissue imaging by imaging MS. Proteomics. 2009; 9(10): 2622-2633.
7. van Hove ER, Smith DF, Fornai L, *et al.* An alternative paper based tissue washing method for mass spectrometry imaging: localized washing and fragile tissue analysis. J Am Soc Mass Spectrom. 2011; 22(10): 1885-1890.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

NON INVASIVE PRENATAL DIAGNOSIS OF ANEUPLOIDY: NEXT GENERATION SEQUENCING OR FETAL DNA ENRICHMENT?

Webb A, Madgett TE, Miran T, Sillence K, Kaushik N, Kiernan M, Avent ND*

***Corresponding Author:** Professor Neil D. Avent, School of Biomedical and Biological Sciences, Faculty of Science and Technology, A411 Portland Square, Drake Circus, Plymouth, Devon, PL4 8AA, UK; Tel.: +44-(0)1752-584884; Fax: +44-(0)1752-584605; E-mail: neil.avent@plymouth.ac.uk

ABSTRACT

Current invasive procedures [amniocentesis and chorionic villus sampling (CVS)] pose a risk to mother and fetus and such diagnostic procedures are available only to high risk pregnancies limiting aneuploidy detection rate. This review seeks to highlight the necessity of investing in non invasive prenatal diagnosis (NIPD) and how NIPD would improve patient safety and detection rate as well as allowing detection earlier in pregnancy.

Non invasive prenatal diagnosis can take either a proteomics approach or nucleic acid-based approach; this review focuses on the latter. Since the discovery of cell free fetal DNA (cffDNA) and fetal RNA in maternal plasma, procedures have been developed for detection for monogenic traits and for some have become well established (*e.g.*, RHD blood group status). However, NIPD of aneuploidies remains technically challenging. This review examines currently published literature evaluating techniques and approaches that have been suggested and developed for aneuploidy detection, highlighting their advantages and limitations and areas for further research.

Keywords: Aneuploidy; Cell free fetal DNA (cff DNA); Non invasive prenatal diagnosis (NIPD).

School of Biomedical and Biological Sciences, University of Plymouth, Plymouth, Devon, UK

INTRODUCTION

Current invasive diagnostic techniques pose a risk to mother and fetus. The National Health Service (NHS) offers prenatal screening to all pregnant women in England [1]. At 11 to 13 weeks gestation a combination of tests (termed the combined test) are performed to screen for abnormalities and score the risk of the fetus having Down's syndrome (trisomy 21) (T21). The combined test includes ultrasound scans to check nuchal translucency and analyzing maternal blood samples for free β human chorionic gonadotrophin and pregnancy-associated plasma protein A concentrations. These are combined with factors such as maternal age, pregnancy histories and familial genetic conditions [1]. Women with high risk pregnancies are then offered prenatal diagnosis. Current diagnostic procedures take samples for karyotyping through invasive means, posing a risk to mother and fetus.

Chorionic villus sampling (CVS) can be performed from 10 weeks; taking a tissue sample from the placenta through the cervix or abdomen. Amniocentesis can be performed from 15 weeks by obtaining a sample of amniotic fluid by passing a needle through the mother's abdomen into the uterus [1]. The rate of miscarriage associated with CVS and amniocentesis is 1.0 to 2.0%. Despite the risks, on average 5.0-10.0% of pregnant women chose to undergo these tests [2], averaging approximately

23,700 invasive diagnostic tests per year [1]. The combined test has a 5.0-9.0% false positive rate [3,4]; consequently mothers with healthy fetuses may choose to undergo unnecessary invasive diagnostic procedures, putting the fetus at risk of iatrogenic spontaneous abortion. Non invasive prenatal diagnosis (NIPD) would pose no such risk.

Increasing maternal age increases risk of Down's syndrome, Edwards' syndrome (trisomy 18) (T18) and Patau's syndrome (trisomy 13) (T13), the three most common autosomal aneuploidies in live births. In the early 1970s, about 5.0% of pregnant women were 35 years or over. However, almost 20.0% are now over 35, increasing the incidence of aneuploid fetuses [5]. Trisomy 18 and T13 are currently not screened for until fetal anomaly ultrasound screening at 18 to 20^{+6} weeks [1]. However, the NHS state that although abortion is legal up to 24 weeks, it should be carried out as early as possible, ideally before 12 weeks [6]. Therefore, preferably, detection of all aneuploidies would be made within the first trimester, which would give greater scope for parental choice. It is hoped that NIPD could help achieve this.

Despite increasing maternal age, approximately 80.0% of Down's syndrome patients are stillborn to those under 35 [7]. Currently diagnostic tests are only offered to those with high risk pregnancies (such as those with abnormal serum protein levels) as the risk of having a fetus with an abnormality must be balanced against the risk of miscarriage. This inevitably results in some abnormalities being missed as women under 35 years are not considered at risk of having a Down's syndrome fetus [7]. Non invasive prenatal diagnosis would allow testing of these low-risk pregnancies. A recent European Commission (EC) funded project, SAFE (Special non-invasive advances in fetal and neonatal evaluation) explored a number of new technologies in NIPD [8,9], and this has been extended in a recent project funded by the UK National Institute for Health Research (NIHR), RAPID, and a EC framework 7 program, Eurogentest 2.

In summary, prenatal testing is important as it allows possible health issues of the baby to be identified before birth, allowing arrangements for immediate care to be made. It also enables parents to make an informed choice regarding whether to terminate the pregnancy. Replacing current invasive tests with NIPD would reduce risk and increase detection rate.

Methods for Non Invasive Prenatal Diagnosis of Aneuploidies. The importance of prenatal diagnosis and risk posed by invasive techniques makes NIPD research a worthwhile commitment morally and commercially. Research originally focused on fetal nucleated cells found in maternal circulation [10-12]. A number of cell types were investigated but subsequently ruled out for a variety of reasons. These include lack of fetal specific markers [13], entrapment in the maternal lungs, confined placental mosaicism [14] and persistence in the maternal circulation years after pregnancy [15]. Furthermore, only one or two fetal nucleated cells are found per mL of maternal blood, further hampering their use [7]. Research now focuses on cell free fetal DNA (cffDNA), which was discovered in maternal plasma in 1997 [10]. It has since been reported to be present from as early as 4 weeks gestation [16], making it possible for NIPD to be achievable earlier in pregnancy than invasive methods. It has a short circulation half-life (≈16 min.) and is undetectable in the maternal circulation within 2 hours postpartum [17], making it specific to the current pregnancy. Technical challenges stem from the fact cffDNA constitutes only 3.0 to 6.0% of cell free DNA in maternal plasma [7,10]. Detection of paternally inherited alleles in maternal plasma is qualitative; however, direct detection of aneuploidies is reliant on dose, therefore quantitative. Therefore, it was originally assumed that direct measurement of fetal chromosome dose in maternal plasma would show maternal chromosome dose and fetal chromosome dose would be lost in the background of maternal DNA. Nevertheless, chromosome dose methods have been developed and are discussed later. Methods for aneuploidy detection originally focused on fetal markers for allelic ratio analysis.

Non Invasive Prenatal Diagnosis in Routine Clinical Practice 2012. Following the discovery of cffDNA in maternal plasma in 1997 [10], testing for fetal RHD blood group status rapidly evolved from risky DNA testing of amniotic fluid samples provided from Liley curve investigations [18] to the world's first routine application of a non invasive prenatal test clinically [19]. Fetal blood grouping using maternal plasma as a source of fetal DNA is now used extensively worldwide [20]. In addition,

the detection of other paternally inherited alleles or chromosome is possible, for example, fetal sexing (for review, see [21]). Most clinically applicable methods utilize the simple real time polymerase chain reaction (ReTi-PCR) approach, however, methods to detect the most commonly requested prenatal diagnostic test, for aneuploidy, require more sophisticated approaches. This is primarily due to the fact that maternal DNA is an admix of fetal (derived from the placenta) and maternal DNA.

METHODS TO DETECT FETAL ANEUPLOIDY BY NON INVASIVE PRENATAL DIAGNOSIS

Methods Based on Allelic Ratio: RNA-Single Nucleotide Polymorphism Allelic Ratio Approach. The presence of free fetal RNA (ffRNA) in maternal plasma was established in 2000 by Poon *et al.* [22] who showed that Y chromosome-specific zinc finger protein (*ZFY*) mRNA could be found in the plasma of women carrying a male fetus. Further studies showed that ffRNA is surprisingly stable [23] and present in maternal plasma as early as 4 weeks gestation [24]. As different tissues express their own individual mRNA profiles, it follows that some mRNAs may be placenta-specific and therefore fetal specific. Ng *et al.* [25] showed that placenta specific mRNAs could be detected in maternal plasma, and in 2007, a placental-specific mRNA transcribed from *PLAC4* on chromosome 21 was discovered [26]; use of these could circumvent the problems caused by maternal background DNA.

Lo *et al.* [26] developed the RNA-SNP (single nucleotide polymorphism) allelic ratio approach. This technique exploits SNPs that cause sequence variation between alleles. A heterozygous euploid fetus should yield equal proportions of each allele, giving an allelic ratio of 1:1. However, a heterozygous triploid fetus would yield the allelic ratio 1:2 or 2:1 [27]. Lo *et al.* [26] used reverse transcription (RT)-PCR to amplify a *PLAC4* SNP containing locus, followed by primer extension. Mass spectrometry was then used to determine allelic ratio (Figure 1). Maternal plasma of 57 women carrying euploid fetuses and 10 carrying T21 fetuses was analyzed. The method had a sensitivity of 90.0% and specificity of 96.0% [26]. Larger trials to refine the normal reference ranges could potentially improve sensitivity and specificity [27]. Regrettably, RNA-SNP allelic ratio methods are limited to fetuses heterozygous for the SNP under analysis. Among 119 placentas genotyped by Lo *et al.* [26] for the most common *PLAC4* SNP (rs8130833), 42.0% were homozygous and the method would have therefore been uninformative.

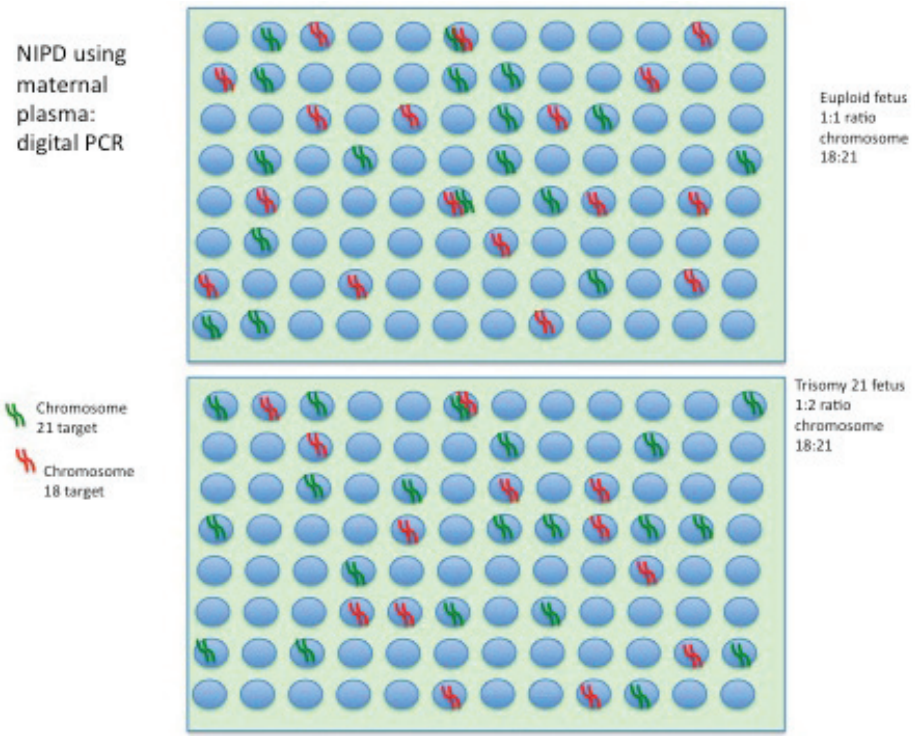


Figure 1. Schematic representation of Digital PCR.

The Digital RNA-Single Nucleotide Polymorphism Allelic Ratio Approach. The digital RNA-SNP allelic ratio approach is an adaptation to the RNA-SNP allelic ratio method made to utilize the sensitivity of digital PCR [28]. Rather than one reaction mix, digital PCR involves thousands of reactions running in parallel. The template strand solution is diluted so that a maximum of one template strand is added to each reaction. Lo *et al.* [28] used 384-well plates for digital ReTi-PCR of the *PLAC4* SNP, rs8130833. Uninformative wells (*i.e.,* those containing no or multiple signals) were discounted and wells containing A or G *PLAC4* allele were counted and their ratio calculated. Using maternal plasma RNA samples, four T21 fetuses and nine euploid were correctly identified. However, although digital PCR is more sensitive than ReTi-PCR, the digital RNA-SNP allelic ratio approach it is still limited to heterozygous fetuses. The methodology entered commercial trials under the governance of Sequenome Inc. (San Francisco, CA, USA) but was subsequently found to be an unreliable technology and unlikely to see routine application in NIPD.

Epigenetic Allelic Ratio Approach. The epigenetic allelic ratio (EAR) approach is similar to the RNA-SNP allelic ratio approach but rather than targeting fetal specific mRNA, it exploits epigenetic phenomena that alter DNA expression without altering sequence; the most well known of these is cytosine methylation. Methylation patterns differ between tissues; genes that are differentially methylated between mother and fetus have been identified allowing an opportunity to selectively target fetal-specific DNA with the use of methylation specific primers [22,27].

Chim *et al.* [29] demonstrated that the *maspin* gene (*SERPINB5*) promoter is unmethylated (U-*mapsin*) in the placenta but hypermethylated in maternal blood cells. *SERPINB5* is located at 18q21.33 [30], providing the opportunity for an EAR-based approach to T18 detection. Tong *et al.* [30] used the EAR approach, first using bisulphite conversion followed by methylation-specific PCR and primer extension reactions designed to distinguish the A and C allele of U-*mapsin* based on size [30]. The test had a sensitivity of 100.0% but had a 9.7% false positive rate [30]. Theoretical modelling suggested that 200 sequence copies were needed at the start of PCR to achieve diagnostic sensitivity and specificity of 97.0%. However, bisulphite conversion has been reported to cause DNA degradation of up to 96.0% [30]. Taking this into account, Tong *et al.* [30] predicted that only 20 sequence copies would have remained in their samples after bisulphite conversion due to low blood volume used; this could explain the high false positive rate.

Although the EAR approach was successfully demonstrated, strategies need to be developed to overcome DNA degradation by bisulphite conversion. Weber *et al.* [31] described a way of enriching methylated DNA using immunoprecipitation with the use of an antibody against 5-methylcytosine; methylated DNA immunoprecipitation (MeDiP) has since been used to successfully enrich fetal hypermethylated DNA [32]. Alternatively, degradation of maternal unmethylated DNA by bisulphite conversion would be acceptable so targeting genes hypermethylated in the placenta but hypomethylated in maternal tissues could be a solution [3].

Despite various groups showing methods based on allelic ratio to work successfully, all are limited to heterozygous fetuses. However, an allelic ratio method could still have potential for use in practice if multiple SNPs, with a combined heterozygosity rate high enough to cover the general population, were analyzed [27]. Allelic ratio methods also assume both alleles are transcribed at an equal rate, which is not always true [4], meaning analysis of transcript ratios could be unreliable. However, successful demonstration of allelic ratio methods suggests this is not problematic. Finally, although allelic ratios may successfully identify trisomies they would not detect monosomies.

Methods Based on Relative Chromosome Dosage. Using microfluidics digital PCR, Lun *et al.* [33] re-evaluated the percentage of cell free DNA that was fetal in maternal plasma and found there was a median 9.7% in the first trimester rising to 20.4% by the third trimester. These levels are higher than previously thought, suggesting direct measurement of fetal chromosome dose may be possible without prior enrichment. Relative chromosome dose approaches work on the principal that in a normal fetus the ratio of two chromosomes should be 2:2. Comparing an affected chromosome in a triploid fetus would give the ratio 3:2. A number of methods based on chromosome dose have been reported and are described below.

Epigenetic-Genetic and Epigenetic-Epigenetic Chromosome Dosage Approaches. Tong *et al*. [34] combined a methylation restricted digest to digest hypomethylated DNA and microfluidics digital PCR to measure *HLCS* on chromosome 21 (a hypermethylated fetal marker), *RASSF1A* on chromosome 3 (a hypermethylated fetal marker) and *ZFY* on the Y chromosome. The epigenetic-epigenetic chromosome dosage approach comparing hypermethylated *HLCS* and *RASSF1A* ratio showed too great an overlap between euploid and T21 fetuses. However the epigenetic-genetic chromosome dosage approach comparing hypermethylated *HLCS* and *ZFY* ratio in euploid and T21 fetuses was discriminative but limited to male fetuses. Other groups have also successfully used the EGG chromosome dose approach, for example, Tsui *et al.* [35] identified T18 fetuses with a sensitivity of 88.9% and specificity of 96.3%. However, Tsui *et al.* [35] also used *ZFY* limiting use to male fetuses. *ZFY* is often used for proof of principal studies but epigenetic-genetic (EGG) chromosome dosage needs to be successfully demonstrated with sex-independent markers before it can be used in practice.

Digital Relative Chromosome Dosage Method. There is a 1.5-fold increase in relative chromosome dose in trisomies. Due to the exponential increase in template strand number in ReTi-PCR, it is not suited to detecting differences in template strand concentrations below 2-fold [4,36,37]. Digital PCR provides a more accurate measure of template strand concentration.

Methodology involves the amplification of multiple targets on different chromosomes in many thousands of different PCR reactions (microfluidic PCR or droplet PCR). In this example, two regions of the human genome have been amplified, chromosomes 18 and 21. The relative abundance of the target DNA is defined directly against one another, and thus a statistical 1:1 ratio of each chromosome would indicate a euploid fetus and a (theoretical) 2:1 ratio (21:18) would indicate aneuploidy. Despite the admix of fetal and maternal DNA in maternal plasma, like next generation sequencing (NGS) (see Figure 2), digital PCR is able to differentiate such a difference, although clearly the actual difference in ratio would not be 2:1 due to the presence of maternal DNA.

The digital relative chromosome dosage (RCD) method directly targets a non polymorphic locus on the chromosome of interest and on a reference chromosome without differentiating between maternal and fetal DNA. Lo *et al.* [28] used digital PCR to measure a non polymorphic locus on chromosome 21 and chromosome 1. The number of wells in which there was a positive PCR for either locus was counted and their relative dose calculated. Euploid and T21 fetuses in artificial mixtures of placental and maternal DNA were successfully identified but the method was not tested using maternal plasma from T21 pregnancies. However, despite RCD not differentiating between maternal and fetal DNA, computer modeling estimated 97.0% of fetuses would be correctly identified by 7680 digital PCRs in cases where the cffDNA constituted 25.0% of the DNA [28]. Nevertheless, this is higher than previously reported levels in maternal plasma. Moreover, other groups have shown the digital RCD approach to be feasible with fetal fractions of 10.0% [38] although some level of enrichment may still first be needed for use in the first trimester.

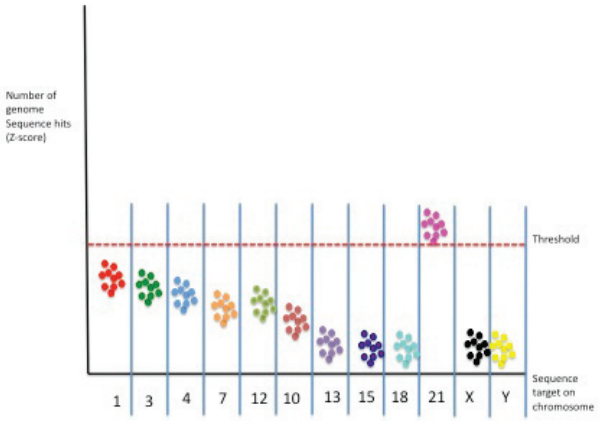


Figure 2. Schematic representation of next-generation sequencing based approaches for NIPD of fetal aneuploidy. The figure represents the numbers of targets from each chromosome (X-axis) identified by NGS as colored dots. Based on statistical counting of each chromosome (Z-score), a higher proportion of chromosome 21 signals can be shown in maternal plasma samples despite this being an admix of maternal and fetal DNA. The approach can include whole genome sequencing, or targeted sequencing of various chromosomes following amplification

Next Generation Sequencing (NGS) or Massively Parallel Sequencing. Massively parallel sequencing (MPS) is a method through which the entire genome can be sequenced using millions of short sequence reads. These are then reassembled by a computer program using genomic databases to compare the reads to the known genomic sequence [3,4,39]. A number of groups have demonstrated

that if cffDNA is sequenced in this way and the sequence reads assigned to each chromosome are then counted, whether a chromosome is over- or under-represented can then be calculated [39,40]. This is similar to the approach taken for the digital relative chromosome dose method. However, NGS generates far more sequence reads, possibly over 60,000 for chromosome 21 alone, suggesting it should be more sensitive to small changes in genomic representation [4]. The NGS could therefore have the potential of identifying partial trisomies, although further research is needed to investigate this.

Fan *et al.* [39] were able to identify nine T21 cases, two T18 and one T13, distinguishing them from six normal euploid cases, using shot gun sequencing (Figure 2). However, the blood samples used in this study were drawn 15 to 30 min. following an invasive procedure, this may have influenced the results by introducing additional fetal DNA into the maternal circulation [3]. Furthermore, Fan *et al.* [39] argue that using a digital PCR assay, they estimated fetal DNA in their samples constituted <10.0% of total cell free DNA, in line with previous studies.

More recent studies have taken into account that differing GC contents between chromosomes results in nonuniform sequencing by MPS and corrected genomic representation to account for CG content. By doing this, Chen *et al.* [40] found that a specificity and sensitivity of 98.9 and 100.0%, respectively, could be achieved for T13 detection and 98.0 and 91.9%, respectively for T18 detection.

In light of recently successful studies, Ehrich *et al.* [41] conducted a blinded study on 449 plasma samples from pregnant women (39 with T21 fetuses), using the most up-to-date sequencing technology available at the time, to investigate NGS for T21 diagnosis. Z-scores described in [42] were used to standardize genomic representation and classify the fetus as euploid or T21. The results were very successful with a sensitivity of 100.0% and specificity of 99.7%, suggesting that if the cost of MPS could be reduced it could be used in practice. Recently, using the single molecule sequencing technology of Helicos (Cambridge, MA, USA) that does not require a prior DNA amplification step, van den Oever *et al.* [43] have demonstrated a greater sensitivity of this platform above that of Illumina GA-II (San Diego, CA, USA) used in previous studies. This may lead to utilization of NGS much earlier in NIPD, perhaps within the first trimester [44], as the study demonstrated detection of trisomy 21.

All three relative chromosome dose methods discussed are polymorphism-independent and have successfully demonstrated the ability to detect trisomies. Of the three relative chromosome dose methods, NGS seems the most promising as, due to the large amount of data it produces, it has high sensitivity and specificity. It can also simultaneously provide information on chromosome dose for all chromosomes, and theoretically, has the potential to detect partial trisomies and monosomies, although this needs to be validated by further research.

Methods to Enrich Fetal DNA. One of the main hindrances on NIPD is the dilution of fetal DNA in maternal blood; this makes the quantitative nature of aneuploidy diagnosis difficult. Therefore, a number of methods have been investigated to enrich or prevent the dilution of fetal DNA in maternal plasma.

Use of Formaldehyde. In 2004, Dhallan *et al*. [45] hypothesized that a significant portion of maternal DNA in maternal plasma is leaked from maternal leukocytes following venipuncture due to physical forces put on them during collection and subsequent handling. Blood was collected from 69 pregnant women in tubes containing a 4.0% formaldehyde neutralizing buffer. Analysis showed the majority of samples had >25.0% fetal DNA and 27.5% had >50.0%, suggesting formaldehyde treatment could successfully prevented leukocyte rupture. However, no untreated blood was collected in this trial so no comparisons between treated and untreated samples could be made. Since 2004, a number of groups have attempted to replicate the enrichment of fetal DNA in maternal plasma using formaldehyde without success [46,47].

Separation Based on Size by Gel Electrophoresis. Another approach to fetal enrichment is to separate cffDNA from maternal based on size. This can be done by gel electrophoresis [48]. In early pregnancy (13 + 2 weeks gestational ages), Li *et al.* [48] found that 85.5% of fetal DNA is less than 0.3 kb. Fetal DNA also constituted 28.4% of the <0.3 kb fraction in maternal plasma, increasing to 68.7% in the third trimester. A study by Chan *et al.* [49] supports Li *et al.* [48], concluding that >99.0% of fetal derived DNA is shorter than 313 bp. Gel electrophoresis has since been used to successfully enrich fetal DNA for detection of point mutation in

β-thalassemia [50]. However, isolation of size fractions by gel electrophoresis is considered too time consuming and prone to contamination to allow its widespread use [3].

Co-Amplification at Lower Denaturation Temperature Polymerase Chain Reaction. In 2008, Li *et al.* [51] reported on (co-amplification at lower denaturation temperature) COLD-PCR, a variation of PCR that can selectively amplify minority alleles from a background of wild-type alleles. This technique works on the basis that even a single nucleotide difference between the minority and wild-type allele may lower the critical denaturation temperature (T_c). If so, the minority allele could be denatured at a lower temperature than the wild-type sequence, allowing only the minority allele sequence to bind with the primers and be amplified [52]. Li *et al.* [51] used COLD-PCR to identify mutations in a number of genes associated with human cancer that had previously been missed and suggested that COLD-PCR could also be used for detection of "fetal alleles in maternal blood." However, in aneuploidies there may be no sequence difference between the maternal and fetal allele of interest. Moreover, it is possible that the shorter length of cffDNA in comparison with maternal cell free DNA may allow COLD-PCR to denature cffDNA at a temperature at which maternal DNA would remain double-stranded, allowing only cffDNA to be amplified. There is currently no published study on the enrichment of fetal DNA from maternal plasma using this method but this could be a promising technique. If fetal DNA could be enriched in this way, not only could it aid detection of aneuploidies but also aid detection of monogenic diseases where a disease allele may have been maternally inherited.

Proteomics for New Down's Syndrome Biomarkers (for review, see [53,54]). Despite the advances in molecular counting technologies, especially NGS and digital PCR, better serum screening biomarkers for Down's syndrome and other conditions are still an option, as NGS and digital PCR methodologies are not yet applicable to mass-scale screening. For this reason several large scale studies were launched to analyze the maternal plasma proteome in pregnancies where the mother was carrying a Down's syndrome fetus [53,55-60]. It is clear that a number of overlapping biomarkers have been identified, but the search for a truly differentiating trisomy 21 biomarker using this approach may still be a way off, but may well be worth the effort expended.

CONCLUSIONS

In summary, despite various non invasive methods invented to diagnose aneuploidy, none are yet used in practice. Methods described in the current literature can be split into methods based on allelic ratio or chromosome dose. Allelic ratio methods are held back by their restriction to heterozygous fetuses, yet if a combination of SNPs could be used, their combined heterozygosity rate could potentially cover the general population. However, allelic ratio approaches would still be unable to diagnose monosomies. Chromosome dose methods have the advantage of being polymorphism-independent and have the potential to diagnose trisomies, monosomies and in the case of NGS-based methods, possibly even partial trisomies but this needs to be verified by further research. For these methods to be introduced into practice, further research needs to be aimed at simplifying the methods and reducing the cost. Another approach would be to selectively enrich cffDNA or deplete maternally derived cell free DNA to overcome the problem of low cffDNA concentration in maternal plasma; this would allow more straight-forward analysis methods to be used. Co-amplification at lower denaturation temperature-PCR seems the most promising enrichment technique discussed but requires further investigation.

With the progress made in NIPD since the discovery of cffDNA, and invention of new sequencing techniques, it is not over optimistic to predict that NIPD will be used in practice in the near future; improving detection rate, allowing earlier diagnosis and eliminating iatrogenic fetal loss and risk to the mother due to invasive procedures.

REFERENCES

1. NHS FASP, 2010. NHS Fetal Anomaly Screening Programme Annual Report 2009-2010. Available at [http://fetalanomaly.screening.nhs.uk/getdata.php?id=11354] Accessed on January 16 2012.
2. Mujezinovic F, Alfirevic Z. Procedure-related complications of amniocentesis and chorionic villous sampling. Obstet Gynecol. 2007; 110(3): 687-694.
3. Go AT, van Vugt JMG, Oudejans CBM. Non-

invasive aneuploidy detection using free fetal DNA and RNA in maternal plasma: recent progress and future possibilities. Hum Reprod Update. 2011; 17(3): 372-382.
4. Hahn S, Lapaire O, Tercanli S, Kolla V, Hösli I. 2011. Determination of fetal chromosome aberrations from fetal DNA in maternal blood: has the challenge finally been met? Expert Rev Molec Med. 2011; 13; 16: e16. Available at [http://www.ncbi.nlm.nih.gov/pmc/articles/ PMC3087311/pdf/S1462399411001852a.pdf] Accessed on January 16 2012.
5. Nicolaides KH. Screening for fetal aneuploidies at 11 to 13 weeks. Prenat Diagn. 2011; 31(1): 7-15.
6. NHS, 2010. When an abortion should be carried out. Available at [http://www.nhs.uk/Conditions/Abortion/ Pages/When-should-it-be-done.aspx] Accessed on January 8 2012.
7. Bischoff FZ, Sinacori MK, Dang DD, *et al.* Cell-free fetal DNA and intact fetal cells in maternal blood circulation: implications for first and second trimester non-invasive prenatal diagnosis. Hum Reprod Update. 2001; 8(6): 493-500.
8. Chitty LS, van der Schoot CE, Hahn S, Avent ND. SAFE–The special non-invasive advances in fetal and neonatal evaluation network: aims and achievements. Prenat Diagn. 2009; 28(2): 83-88.
9. Maddocks DG, Alberry MS, Attilakos G, *et al.* The SAFE project–towards non-invasive prenatal diagnosis. Biochem Soc Transactions. 2009; 37: 460-465.
10. Lo YMD, Corbetta N, Chamberlain PF, *et al.* Presence of fetal DNA in maternal plasma and serum. Lancet 1997; 350(9076): 485-487.
11. Lo YMD. Non-invasive prenatal diagnosis using fetal cells in maternal blood. J Clin Pathol. 1994; 47(12): 1060-1065.
12. Herzenburg LA, Bianchi DW, Schroder J, Cann HM, Iverson GM. Fetal cells in the blood of pregnant women: Detection and enrichment by fluorescence-activated cell sorting. Proc Natl Acad Sci USA. 1979; 76(3): 1453-1455.
13. Avent ND, Plummer ZE, Madgett TE, *et al.* Post-genomics studies and their application to non-invasive prenatal diagnosis. Semin Fetal Neonat Med. 2008; 13: 91-98.
14. Wapner RJ. Invasive prenatal diagnostic techniques. Semin Perinatol. 2005; 29(6): 401-404.
15. Bianchi DW, Zickwolf GK, Weil GJ, Sylvester S, DeMaria MA. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. Proc Natl Acad Sci USA. 1996; 93(2): 705-708.
16. Illanes S, Denbow M, Kailasam C, Finning K, Soothill PW. Early detection of cell-free fetal DNA in maternal plasma. Early Hum Dev. 2007; 83(9): 563-566.
17. Lo YMD, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. Am J Hum Genet. 1999; 64(1): 218-224.
18. Avent ND, Finning KM, Martin PG, Soothill PW. Prenatal determination of fetal blood group status. Vox Sang. 2000; 78(2): 155-162.
19. Finning KM, Martin PG, Soothill PW, Avent ND. Prediction of fetal D status from maternal plasma: introduction of a new noninvasive fetal RHD genotyping service. Transfusion. 2002; 42(8): 1079-1085.
20. Scheffer PG, van der Schoot CE, Page-Christiaens GCML, de Haas M. Noninvasive fetal blood group genotyping of rhesus D, c, E and of K in alloimmunised pregnant women: evaluation of a 7 year clinical experience; Br J Obstet Gynaecol. 2011; 118(11): 1340-1348.
21. Avent ND, Chitty LS. Non-invasive diagnosis of fetal sex; utilisation of free fetal DNA in maternal plasma and ultrasound. Prenat Diagn. 2006; 26(7): 598-603.
22. Poon LLM, Leung TN, Lau TK, Chow KCK, Lo YMD. Differential DNA methylation between fetus and mother as a strategy for detecting fetal DNA in maternal plasma. Clin Chem. 2002; 48(1): 35-41.
23. Tsui NBY, Ng EKO, Lo YMD. Stability of endogenous and added RNA in blood specimens, serum, and plasma. Clin Chem. 2002; 48(10): 1647-1653.
24. Chiu RW, Lui WB, Cheung MC. *et al.* Time profile of appearance and disappearance of circulating placenta-derived mRNA in maternal plasma. Clin Chem. 2006; 52(2): 313-316.
25. Ng EKO, Tsui NBY, Lau TK, *et al.* mRNA of placental origin is readily detectable in maternal plasma. Proc Natl Acad Sci USA. 2003; 100(8): 4748-4753.
26. Lo YMD, Tsui NBY, Chiu RWK, *et al.* Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. Nat Med. 2007; 13(2): 218-223.
27. Lo YMD. Noninvasive prenatal detection of fetal chromosomal aneuploidies by maternal plasma nucleic acid analysis: a review of the current state of the art. Br J Obstet Gynaecol. 2008; 116(2): 152-157.
28. Lo YMD, Lun FMF, Chan KCA, *et al.* Digital PCR for the molecular detection of fetal chromosomal aneuploidy. Proc Natl Acad Sci USA. 2007; 104(32): 13116-13121.
29. Chim SSC, Tong YK, Chiu RWK, *et al*. Detection of the placental epigenetic signature of the maspin gene in maternal plasma. Proc Natl Acad Sci USA. 2005; 102(41): 14753-14758.
30. Tong YK, Ding C, Chiu RWK, *et al.* Noninvasive prenatal detection of fetal trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: theoretical and empirical considerations. Clin Chem. 2006; 52(12): 2194-2202.

31. Weber M, Davies JJ, Wittig D, *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet. 2005; 37(8): 853-62.
32. Papageorgiou EA, Fiegler H, Rakyan V, *et al.* Sites of Differential DNA Methylation between Placenta and Peripheral Blood: Molecular Markers for Noninvasive Prenatal Diagnosis of Aneuploidies. Am J Pathol. 2009; 174(5): 1609-1618.
33. Lun FMF, Chiu RWK, Chan KCA, *et al.* Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. Clin Chem. 2008; 54(10): 1664-1672.
34. Tong YK, Jin S, Chiu RWK, *et al.* Noninvasive prenatal detection of trisomy 21 by an epigenetic-genetic chromosome-dosage approach. Clin Chem. 2010; 56(1): 90-98.
35. Tsui DWY, Lam YMD, Lee WS, *et al.* Systematic identification of placental epigenetic signatures for the noninvasive prenatal detection of Edward's syndrome. PLoS One 2010; 5(11): e15069.
36. Vogelstein B, Kinzler KW. 1999. Digital PCR. Proc Natl Acad Sci USA. 1999; 96(16): 9236-9241.
37. Zimmermann BG, Grill S, Holzgreve W, *et al.* Digital PCR: a powerful new tool for noninvasive prenatal diagnosis? Prenat Diagn. 2008; 28(12): 1087-1093.
38. Fan HC, Quake SR. Detection of aneuploidy with digital polymerase chain reaction. Anal Chem. 2007; 79(19): 7576-7579.
39. Fan HC, Blumenfeld YJ, Chitkara U, *et al.* Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proc Natl Acad Sci USA. 2008; 105(42): 16266-16271.
40. Chen EZ, Chiu RWK, Sun H, *et al.* Noninvasive prenatal diagnosis of fetal trisomy 18 and trisomy 13 by maternal plasma DNA sequencing. PLoS One 2011; 6(7): e21791.
41. Ehrich M, Deciu C, Zwiefelhofer T, *et al.* Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. Am J Obstet Gynecol. 2011; 204(3): 205.e1-11.
42. Chiu RWK, Chan KCA, Gao Y, *et al.* Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc Natl Acad Sci USA. 2008; 105(51): 20458-20463.
43. van den Oever JM, Balkassmi S, Verweij EJ, *et al.* Single molecule sequencing of free DNA from maternal plasma for noninvasive trisomy 21 detection. Clin Chem. 2012; 58(4): 699-706.
44. Avent ND. Refining noninvasive prenatal diagnosis with single-molecule next generation sequencing. Clin Chem. 2012; 58(4): 657-658.
45. Dhallan R, Au WC, Mattagajasingh S, *et al.* Methods to increase the percentage of free fetal DNA recovered from the maternal circulation. JAMA. 2004; 291(9): 1114-1119.
46. Chung GTY, Chiu RWK, Chan KCA, *et al.* Lack of dramatic enrichment of fetal DNA in maternal plasma by formaldehyde treatment. Clin Chem. 2005; 51(3): 655-658.
47. Chinnapapagari SKR, Holzgreve W, Lapaire O, *et al.* Treatment of maternal blood samples with formaldehyde does not alter the proportion of circulatory fetal nucleic acid (DNA and RNA) in maternal plasma. Clin Chem. 2005; 51(3): 652-655.
48. Li Y, Zimmermann B, Rusterholz C, *et al.* Size separation of circulatory DNA in maternal plasma permits ready detection of fetal DNA polymorphisms. Clin Chem. 2004; 50(6): 1002-1011.
49. Chan KCA, Zhang J, Hui ABY, *et al.* Size distributions of maternal and fetal DNA in maternal plasma. Clin Chem. 2004; 50(1): 88-92.
50. Li Y, Di Naro E, Vitucci A, *et al.* Detection of paternally inherited fetal point mutations for β-thalassemia using size-fractionated cell-free DNA in maternal plasma. JAMA. 2005; 293(7): 843-849.
51. Li J, Wang L, Mamon H, *et al.* Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. Nat Med. 2008; 14(5): 579-584.
52. Li J, Makrigiorgos GM. COLD-PCR: a new platform for highly improved mutation detection in cancer and genetic testing. Biochem Soc Transactions. 2009; 37(2): 427-432.
53. Kolialexi A, Tsangaris GT, Papantoniou N, *et al.* Application of proteomics for the identification of differentially expressed protein markers for Down syndrome in maternal plasma. Prenat Diagn. 2008; 28(8): 691-698.
54. Choolani M, Narasimkhan K, Kolla V, Hahn S. Proteomic technologies for prenatal diagnostics: advances and challenges ahead. Expert Rev Proteomics. 2009; 6: 87-101.
55. Tsangaris GT, Karamessinis P, Kolialexi A, *et al.* Proteomic analysis of amniotic fluid in pregnancies with Down syndrome. Proteomics. 2006; 6(15): 4410-4419.
56. Nagalla SR, Canick JA, Jacob T, *et al.* Proteomic analysis of maternal serum in Down syndrome: identification of novel protein biomarkers. J Proteome Res. 2007; 6: 1245-1257.
57. Kolla V, Jenö P, Moes S, et al. Quantitative proteomics analysis of maternal plasma in Down syndrome pregnancies using isobaric tagging reagent (iTRAQ). J Biomed Biotechnol. 2010; 2010: 952047 [Epub 2009 Nov 5].

58. Heywood WE, Madgett TE, Wang D, *et al.* 2D-DIGE analysis of maternal plasma for potential biomarkers for Down syndrome. Proteome Sci. 2011; 9: 56.
59. Heywood WE, Mills K, Wang D, *et al.* Identification of new biomarkers for Down's syndrome in maternal plasma. J Proteomics. 2012; in press.
60. Heywood WE, Wang D, Madgett TE, *et al.* The development of a peptide -SRM based tandem mass spectrometry assay for prenatal screening of Down syndrome. J Proteomics. 2012; in press.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

PROTEOMICS OF THE SPERMATOZOON

Oliva R[1,*], Ballescá JL[2]

***Corresponding Author:** Professor Rafael Oliva, Laboratorio de Genética Humana, Facultad de Medicina, Universitat de Barcelona, Hospital Clinic and Institut d'Investigacions Biomèdiques August Pi I Sunyer (DIBAPS), C/Casanova 143, 08036 Barcelona, Spain; Tel.: +34934021877; Fax:+34934035278; E-mail: roliva@ub.edu

ABSTRACT

The study of the sperm proteins is crucial for understanding its normal function and alterations in infertile patients. The sperm is a highly specialized cell with a very large flagella, with little cytoplasm and a highly condensed nucleus. The most abundant proteins in the nucleus of mammalian sperm are the protamines. The main functions of the protamines are the condensation of the DNA, possibly contributing to the generation of a more hydrodynamic sperm head and to the protection of the genetic message. However, in addition to protamines, about 5.0-15.0% of the paternal genome is also complexed with histones and histone variants. It has also demonstrated a differential distribution of genes in regions associated with histone and protamine-associated regions, suggesting a potential epigenetic relevance in embryonic development. More recently, detailed lists of proteins have been described corresponding to the different compartments of the sperm cell thanks to the application of recent proteomic techniques based on mass spectrometry (MS). Differential proteomics is also being applied to identify the presence of protein abnormalities found in infertile patients.

Keywords: Proteomics, Proteome, Sperm chromatin, Epigenetics, Infertile

[1] Servicio de Bioquímica y Genética Molecular, Centre de Diagnòstic Biomèdic, Hospital Clínic de Barcelona, y Laboratorio de Genética Humana, Facultad de Medicina, Universidad de Barcelona, Hospital Clinic and Institut d'Investigacions Biomèdiques August Pi I Sunyer (DIBAPS), Barcelona, Spain

[2] Institut Clínic de Ginecología Obstetricia i Ginecologia, Hospital Universitari Clínic de Barcelona, Barcelona, Spain

The Nucleohistone-Nucleoprotamine Transition and Organization of the DNA in the Sperm Nucleus. Spermatogenesis involves radical changes in chromatin structure to give rise to the mature sperm [1,2]. The nucleosome structure present in spermatogonia, spermatocytes and round spermatids, is disassembled in spermiogenesis and is temporarily replaced by transition proteins and finally by protamines [1-4].

While most of the genome in the sperm cell (about 85.0-95.0%) is tightly packaged by protamines in the form of toroidal structures, it is also important to note that about 5.0-15.0% of sperm DNA is organized by histone proteins, many of which are sperm-specific variants [3-5]. The distribution of genes in genomic regions organized by protamine and the genomic regions organized by histones is not random. Recent studies based on analysis of the paternal genome associated with each of these domains using microarrays, have led to the basic conclusion that the regions associated

with the nucleohistone are associated with gene regulatory regions [6]. In another recent study [7], based on massive genome sequencing, it was found that nucleosomes associated regions are significantly enriched in genes important for development, including imprinted genes, microRNAs, Hox genes, promoters and transcription developmental genes and signaling factors. It has also been shown that histone modifications (H3K4me2, H3K27me3) are reached at certain loci associated with developmental genes, and promoters associated with developmental genes are hypomethylated in the sperm, but are methylated during maturation [7,8].

In addition to these epigenetic marks determined by the differential distribution of genes in the domains associated with the nucleohistone and nucleoprotamine, other types of epigenetic information are potentially transmitted by the sperm nucleus to the oocyte. One of the best known is contrasted DNA methylation. More recently, the identification of RNAs present in the sperm and the demonstration of oocyte transfer, opens the possibility of their role in fertilization. Another potential source of epigenetic information could be the presence of other proteins in the sperm nucleus, in addition to histones and protamines [9,10].

More recently, proteomic analysis of proteins identified in mature sperm has provided some unexpected results. For example, transcription factors, DNA binding proteins and proteins involved in the metabolism of the chromatin in cells that are transcriptionally inactive [9,10]. The catalogs for the proteomes of human sperm are available [9,11,12]. Most notable is the presence of proteins such as histone acetyltransferase and deacetylase, histone methyltransferase, DNA methyltransferase, topoisomerase, helicase, transcription factors, zinc finguers, homeobox proteins, cromodominio proteins, centrosomal proteins, and telomerase in cells that are transcriptionally inert and have at least 85.0% of their DNA packaged by protamines [9]. A crucial question is whether these transcription factors and proteins newly identified in the cores, represent remnants of the process of spermatogenesis or are making some regions of the paternal genome and have an epigenetic basis [9,13].

Abnormalities in the content of protamine in subfertile patients have already been described over 20 years ago [14]. Subsequently, other studies confirmed the link between abnormal protamine content and alterations in sperm parameters in infertile patients [15,16]. One of the potential causes of abnormal protamine ratio (P1/P2) can be found in abnormal processing of protamine 2 and increased protamine precursors in a subset of patients [15,17]. The results of the content of protamines and histones have been correlated with alterations in the integrity of DNA and the results of assisted reproduction [18].

Identification of Sperm Proteins and Their Alterations in the Spermatozoa Through Proteomic Techniques. Essentially, two different alternatives can be used to study the sperm proteome through mass spectrometry (MS) (Figure 1): ***1)*** two dimensional (2D) separation of the proteins followed by their identification by Matrix-assisted laser desorption/ionization (MALDI)-MS or liquid chro-matography-tandem MS (LC-MS/MS), and ***2)*** the initial digestion of proteins to generate peptides, followed by separation and LC-MS/MS analysis [9]. The first alternative generally involves the separation of proteins using isoelectric focusing and is followed by polyacrylamide gel electrophoresis (PAGE) in the presence of sodium dodecyl sulfate (SDS) for separating proteins in a second dimension based on their molecular weight. This alternative has been widely used in the past to identify many proteins present in the sperm cell [11,19]. Of the two alternatives, the initial generation of peptides and analysis by LC-MS/MS is of much higher throughput. For example, through 2D and MALDI-TOF (time of flight) or LC-MS/ MS, it has been possible to identify some hundreds of proteins [11,20], whereas the generation of peptides followed by LC-MS/MS allows the identification of up to about 1000 different proteins [9,12].

In addition to the generation of catalogs of proteins, proteomics has also been applied to the identification of the presence of anomalies in infertile patients. There are several strategies to analyze the differential protein content in two or more different samples. One method is 2D-DIGE (differential in-gel electrophoresis) and is based on the differential identification of fluorochrome-labeled proteins extracted from the control (for example, labeled green) and experimental cells (for example, labeled red). This is followed by mixing of the proteins and their separation in the same 2D system, followed by detection that can detect increased or decreased pro-

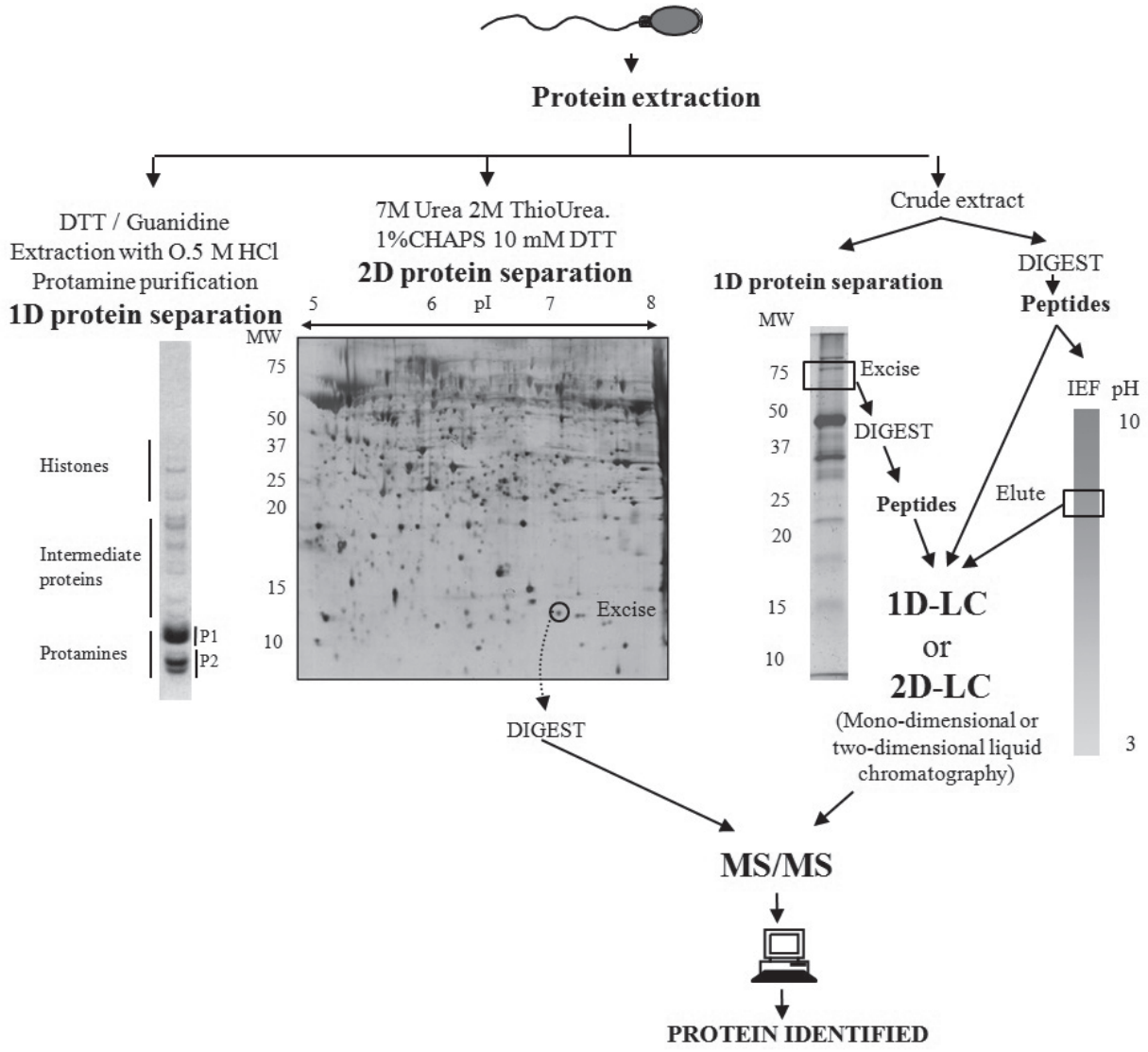


Figure 1. Strategies available for studying the sperm cell proteome. Typical extraction from sperm protamines consisting of reducing the disulfide bridges of protamines with DTT/guanidine hydrochloride, followed by extraction with 0.5 M HCl, precipitation and purification of the proteins and their separation by PAGE acid (left). With this strategy it is possible to identify the protamine 1 (P1) and a set of bands corresponding to the family of protamine 2 (P2). To analyze the total proteome it is possible to use 2D electrophoresis of the proteins. The identification of proteins is then performed *via* MS. A more robust strategy involves the initial generation of peptides followed by their separation by liquid chromatography and identification using MS (right). Based on [9] with modifications.

teins, observing the deviation of the fluorescence to one of the fluorochromes [9,12]. Another alternative is the quantification and comparison of the relative abundance of the different proteins in separate gels. Newer strategies are being developed based on non radioactive isotopic labeling of the test samples and control [9,12].

The first description of the potential of 2D proteomic analysis in the study of defects in sperm was performed in a patient with repeated failure of *in vitro* fertilization techniques [20]. The proteome of this patient showed 20 differences compared with controls, and identified several proteins differentials. It was later applied to the identification of the differential proteins in astenozoospermic patients, oligozoospermic patients, and patients with alterations in the content of protamines or the integrity of DNA [19].

The application of proteomics techniques in andrology and reproductive biology is in its infancy

but the data available to date indicate their enormous potential. It is foreseeable that in the future it will allow the molecular dissection of the various causes of male infertility, allowing both the identification of the pathophysiologic mechanisms involved and its application to the diagnosis, prognosis, and development of new therapeutic strategies.

ACKNOWLEDGMENTS

This study was funded in part by the project of the Ministry of Science and Technology (BFU2009-07118), Spain.

REFERENCES

1. Mezquita C. Chromatin composition, structure and function in spermatogenesis. Revis Biol Celular. 1985; 5: V-XIV, 1-124.
2. Oliva R, Dixon GH. Vertebrate protamine genes and the histone-to-protamine replacement reaction. Prog Nucleic Acid Res Mol Biol. 1991; 40: 25-94.
3. Oliva R. Protamines and male infertility. Hum Reprod Update. 2006; 12(4): 417-435.
4. Balhorn R. The protamine family of sperm nuclear proteins. Genome Biol. 2007; 8(9): 227.1-227.8.
5. Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. Science. 1987; 236(4804): 962-964.
6. Arpanahi A, Brinkworth M, Iles D, *et al.* Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res. 2009; 19(8): 1338-1349.
7. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. Nature. 2009; 460(7254): 473-478.
8. Jodar M, Oriola J, Mestre G, *et al.* Polymorphisms, haplotypes and mutations in the protamine 1 and 2 genes. Int J Androl. 2011; 34(5 Pt. 1): 470-485.
9. Oliva R, de Mateo S, Estanyol JM. Sperm cell proteomics. Proteomics. 2009; 9(4): 1004-1017.
10. de Mateo S, Castillo J, Estanyol JM, Ballescà JL, Oliva R. Proteomic characterization of the human sperm nucleus. Proteomics. 2011; 11(13): 2714-2726.
11. Martinez-Heredia J, Estanyol JM, Ballesca JL, Oliva R. Proteomic identification of human sperm proteins. Proteomics. 2006; 6(15): 4356-4369.
12. Baker MA, Reeves G, Hetherington L, Muller J, Baur I, Aitken RJ. Identification of gene products present in Triton X-100 soluble and insoluble fractions of human spermatozoa lysates using LC-MS/MS analysis. Proteomics Clin Appl. 2007; 1(5): 524-532.
13. Rousseaux S, Reynoird N, Escoffier E, Thevenon J, Caron C, Khochbin S. Epigenetic reprogramming of the male genome during gametogenesis and in the zygote. Reprod Biomed Online. 2008; 16(4): 492-503.
14. Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia. 1988; 44(1): 52-55.
15. de Yebra L, Ballesca JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in humans. J Biol Chem. 1993; 268(14): 10553-10557.
16. Mengual L, Ballescá JL, Ascaso C, Oliva R. Marked differences in protamine content and P1/P2 ratios in sperm cells from percoll fractions between patients and controls. J Androl. 2003; 24(3): 438-447.
17. Bench G, Corzett MH, De Yebra L, Oliva R, Balhorn R. Protein and DNA contents in sperm from an infertile human male possessing protamine defects that vary over time. Mol Reprod Dev. 1998; 50(3): 345-353.
18. Castillo J, Simon L, de Mateo S, Lewis S, Oliva R. Protamine/DNA ratios and DNA damage in native and density gradient centrifuged sperm from infertile patients. J Androl. 2011;32(3): 324-332.
19. Martinez-Heredia J, de Mateo S, Vidal-Taboada JM, Ballesca JL, Oliva R. Identification of proteomic differences in asthenozoospermic sperm samples. Hum Reprod. 2009; 23(4): 783-791.
20. Pixton KL, Deeks ED, Flesch FM, *et al.* Sperm proteome mapping of a patient who experienced failed fertilization at IVF reveals altered expression of at least 20 proteins compared with fertile donors: case report. Hum Reprod. 2004; 19(6): 1438-1447.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

GENETIC CAUSES OF MALE INFERTILITY

Plaseska-Karanfilska D[1,*], Noveski P[1], Plaseski T[2],
Maleva I[1], Madjunkova S[1], Moneva Z[1]

***Corresponding Author:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

INTRODUCTION

Infertility is a major health problem today, affecting about 15.0% of couples trying to have a child. Impaired fertility of the male is causative in 20.0% of infertile couples and contributory in up to another 30.0-40.0%. Infertility already affects about 5.0-7.0% of the general male population and may further increase in the future, considering the apparent trend of declining sperm count in industrialized countries. Despite enormous progress in the understanding of human reproductive physiology, the underlying cause of male infertility remains undefined in about 50.0% of cases, which are referred to as idiopathic infertility [1]. Most idiopathic cases are likely to be of genetic origin because the number of genes involved in human spermatogenesis is probably over 1 thousands. At present, only a few of the genes implicated in the processes of testis determination, testis descent and spermatogenesis have routine clinical importance. These include the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, whose mutations cause cystic fibrosis and absence of vas deferens and the androgen receptor (*AR*) gene, whose mutations cause the androgen insensitivity syndrome and spermatogenic damage.

Common Genetic Causes of Male Infertility. Chromosomal anomalies and microdeletions of the azoospermia factor (AZF) regions of the Y chromosome are the only commonly known genetic causes of spermatogenic failure. The frequency of these two genetic anomalies increases with the severity of the spermatogenic defect, reaching up to an overall 30.0% (15.0% karyotype abnormalities and 15.0% of AZF microdeletions) in azoospermic men.

Sex chromosome aneuploidies, such as 47,XXY (Klinefelter's syndrome), 47,XYY and 46,XX males are the most common chromosome anomalies occurring at birth and in the population of infertile males [2]. Klinefelter's syndrome is a form of primary testicular failure with a high prevalence in infertile men, up to 5.0% in severe oligozoospermia and 10.0% in azoospermia.

Y chromosome microdeletions represent the etiological factor of 10.0-15.0% of idiopathic azoospermia and severe oligozoospermia [3]. The frequency of AZF deletions in infertile men ranges from 5.0 to 20.0% in worldwide surveys [4]. Y chromosome microdeletions are found almost exclusively in patients with azoospermia or severe oligozoospermia [5]. The prevalence of Y chromosome microdeletions in the infertile males from the Republic of Macedonia is 6.4%, in patients with azoospermia 16.7% and 2.8% in those with severe oligozoosper-

[1] Macedonian Academy of Sciences and Arts, Research Center for Genetic Engineering and Biotechnology "Georgi D. Efremov," Skopje, Republic of Macedonia
[2] Clinic of Endocrinology and Metabolic Disorders, Faculty of Medicine, Skopje, Republic of Macedonia

mia [6]. Deletions most frequently involve the AZFc region, less frequently the AZFb region, and only rarely the AZFa region. The most frequent deletions in Macedonian males are AZFc deletions, while AZFa deletions have not been detected [7,8].

Partial deletions within the AZFc region (gr/gr and b2/b3) that remove smaller portions of the AZFc region (1.6 and 1.8 Mb) are much more common and are present at various frequencies in different Y chromosome haplogroups [9]. While the association of the complete AZFc deletion with spermatogenic failure is well established, the role of partial AZFc deletions on spermatogenesis and male infertility is still controversial.

In addition to deletions, different duplications at the AZFc region have been reported. Duplications can occur on a chromosome with a partial AZFc deletion and generate a chromosome with four *DAZ* genes, but lacking some sequence tagged site (STS) markers [10,11]. Recently, an AZFc partial duplication has been shown to be a risk factor for male infertility in Taiwan [12].

Screening for Common Genetic Causes of Male Infertility by Quantitative Fluorescent-Polymerase Chain Reaction. Screening for chromosomal abnormalities is usually done by cytogenetic analysis and for AZF deletions by polymerase chain reaction (PCR) analysis of several STSs in the three AZF regions. Recently, we described a multiplex quantitative fluorescent (QF)-PCR method that allows simultaneous detection of the most common genetic causes of male infertility, *i.e.,* sex chromosomal aneuploidies and AZFc and AZFb deletions, and some potential risk factors such as partial AZFc deletions/duplications and *AR* CAG repeats [8]. This multiplex QF-PCR analysis was shown to be a rapid, simple, reliable and inexpensive method that can be used as a first-step genetic analysis in infertile patients. Recently, we presented a modified system, where we have included additional markers in the AZFa and AZFb region, as well as a marker for determination of the X/chromosome 3 ratio [13].

Our results showed that Klinefelter's syndrome and complete AZFc deletions are the most common genetic causes of azoospermia. Partial AZFc deletions as well as AZFc duplications were present in both infertile and fertile men. They may represent a risk factor for male infertility when present on certain Y chromosomal backgrounds.

Gene Polymorphisms and Male Infertility. Analysis of Y chromosome haplogroups, defined by single nucleotide polymorphisms (SNPs), has become a standard approach for studying the origin of human populations and measuring the variability among them. A few groups have studied the possible association of Y chromosome haplogroups with male infertility and Y chromosome microdeletions, but conflicting results have been published. Some recent studies suggested that a Y chromosome background is an important factor that affects partial AZFc deletion formation and its contribution to spermatogenic failure [14].

We have used a hierarchical analysis of 28 SNP markers by multiplex PCR followed by single base extension reactions using a multiplex SNaPshot kit to determine the Y chromosome haplogroups in men from our country [15]. Our initial results showed slight differences in the distribution of the Y chromosome haplogroups such as higher frequency of the R1a haplogroup in infertile patients with a milder phenotype in comparison with those with azoospermia and severe oligozoospermia and fertile controls.

We have studied in detail the Y chromosomal background of different Y chromosome deletions detected in men from our country. Several different Y chromosome haplogroups were determined in men with complete AZFc (b2/b4) deletions and gr/gr deletions. All infertile males with b2/b3 deletion belong to the Hgr E3b1 anomaly, while the only fertile man with this deletion falls within the Hgr N3 anomaly. Most of the men with the b2/b4 duplication, both infertile and fertile, were identified as Hgr R1a, but the frequency of this Hgr was higher in infertile men. There was also a difference in the distribution of the Y chromosome haplogroups in males with the b2/b3 duplication.

The analysis of polymorphisms in genes involved in spermatogenesis represents one of the most exciting areas of research in genetics of male infertility [16]. Polymorphisms in these genes are considered potential risk factors that may contribute to the severity of spermatogenic failure. Polymorphisms in different genes [CAG repeats in *AR* and DNA polymerase γ (*POLG*) genes, C677T mutation in 5-methylenetetrahydrofolate reductase (*MTHFR*), A260G and A386G in the *DAZL* gene, different polymorphisms in *FSHR*, *ER*α, protamine

1 and 2, *etc.*] have been studied for possible association with male infertility but many of them have presented contradictory results. It is likely that only polymorphisms in association with a specific genetic background and/or with environmental factors can lead to spermatogenic impairment. We have also studied the possible association of several different polymorphisms with male infertility. There was no association between the *POLG* polymorphism and infertility in Macedonian men [17]. We found a significantly higher percentage of long CAG repeats in patients with mild oligozoospermia indicating the possible association of CAG repeat numbers in exon 1 of the *AR* gene and mild oligozoospermia [18]. Our preliminary results suggest that there is no association between the *MTHFR* C677T, *MTHFR* A1298C, *MTR* A2756G and *MTRR* A66G polymorphisms and male infertility. Of the nine SNPs evaluated in eight different genes (*FASLG*, *JMJDIA*, *LOC203413*, *TEX15*, *BRDT*, *OR2W3*, *INSR* and *TAS2R38*), we found significant association for three SNPs (rs5911500 in the *LOC203413*, rs3088232 in the *BRDT* and rs11204546 in the *OR2W3* genes, respectively) [19].

Copy number variations (CNVs) represent an important source of genetic diversity with remarkable differences between individuals. Copy number variations can cause spermatogenic failure by their increased number or specific distribution that could result in defective recombination, meiotic failure and loss of germ cells. Copy number variations might also affect the activity of genes important for spermatogenesis. The first study that investigated CNVs in patients with severe oligozoospermia and Sertoly cell only syndrome (SCOS) was published only recently [20]. This study provided a number of candidate genes, possibly causing or being risk factors for, spermatogenic failure. Using array CGH analysis we have also identified several CNVs (*UGT2B17* gene on chr4 q13.2; *STEAP2* gene on chr7 q21.13; *TPTE* gene on chr21 p11.2-11.1 and *H2BFWT* on chrX q22.2) that might be associated with impaired spermatogenesis and male infertility.

ACKNOWLEDGMENTS

This study was supported by grant CRP/MAC09-01 from the International Centre for Genetics Engineering and Biotechnology, Trieste, Italy (to D. Plaseska-Karanfilska). We are grateful to Dr. Rafael Oliva for the training of Ms. Sanja Kiprijanovska in the genomic and proteomic investigation of spermatozooon that she received during her stay at the University of Barcelona, Faculty of Medicine, Department of Physiological Sciences, Human Genetics Laboratory, Barcelona, Spain. Ms. Kiprijanovska's 2-week stay was supported by the FP7 project No. 229458 from the European Commission.

REFERENCES

1. Ferlin A, Arredi B, Foresta C. Genetic causes of male infertility. Reprod Toxicol. 2006; 22(2): 133-141.
2. Gekas J, Thepot F, Turleau C, *et al.* Chromosomal factors of infertility in candidate couples for ICSI: an equal risk of constitutional aberrations in women and men. Hum Reprod. 2001; 16(1): 82-90.
3. Foresta C, Ferlin A, Moro E, Scandellari C. Y chromosome. Lancet. 2000; 355(9199): 234-235.
4. Krausz C, Forti G, McElreavey K. The Y chromosome and male fertility and infertility. Int J Androl. 2003; 26(2): 70-75.
5. Simoni M, Kamischke A, Nieschlag E. Current status of the molecular diagnosis of Y-chromosomal microdeletions in the work-up of male infertility. Initiative for international quality control. Hum Reprod. 1998; 13(7): 1764-1768.
6. Plaseski T, Dimitrovski C, Kocevska B, Efremov GD, Plaseska-Karanfilska D. The prevalence of Y chromosome microdeletions among infertile males from the Republic of Macedonia. Balkan J Med Genet. 2003; 6(1&2): 39-44.
7. Plaseski T, Novevski P, Kocevska B, Dimitrovski C, Efremov GD, Plaseska-Karanfilska D. AZF deletions in infertile men from the Republic of Macedonia. Prilozi. 2006; 27(1): 5-16.
8. Plaseski T, Noveski P, Trivodalieva S, Efremov GD, Plaseska-Karanfilska D. Quantitative fluorescent-PCR detection of sex chromosome aneuploidies and AZF deletions/duplications. Genet Test. 2008; 12(4): 595-605.
9. Vogt PH. AZF deletions and Y chromosomal haplogroups: history and update based on sequence. Hum Reprod Update. 2005; 11(4): 319-336.
10. Repping S, Skaletsky H, Brown L, *et al.* Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. Nat Genet. 2003; 35(3): 247-251.
11. Repping S, van Daalen SK, Korver CM, *et al.* A family of human Y chromosomes has dispersed through-

out northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. Genomics. 2004; 83(6): 1046-1052.
12. Lin YW, Hsu LC, Kuo PL, *et al.* Partial duplication at AZFc on the Y chromosome is a risk factor for impaired spermatogenesis in Han Chinese in Taiwan. Hum Mutat. 2007; 28(5): 486-494.
13. Plaseska-Karanfilska D, Noveski P, Plaseski T. PNaTP, detection of the most common genetic causes of male infertility by quantitative fluorescent (QF)-PCR analysis. In: Plaseska-Karanfilska, Ed. Genetics of Human Diseases. Rijeka: Croatia. 2011: 203-222.
14. Yang Y, Ma M, Li L, *et al.* Y chromosome haplogroups may confer susceptibility to partial AZFc deletions and deletion effect on spermatogenesis impairment. Hum Reprod. 2008; 23(9): 2167-2172.
15. Noveski P, Trivodalieva S, Efremov GD, Plaseska-Karanfilska D. Y-chromosome single nucleotide polymorphisms typing by SNaPshot minisequencing. Balkan J Med Genet. 2010; 13(1): 9-16.
16. Krausz C, Giachini C. Genetic risk factors in male infertility. Arch Androl. 2007; 53(3): 125-133.
17. Plaseski T, Noveski P, Kocevska B, Dimitrovski C, Efremov GD, Plaseska-Karanfilska D. CAG repeat polymorphism of the mitrochondrial DNA polymerase gamma gene in Macedonian infertile and fertile men. Balkan J Med Genet. 2007; 10(2): 37-41.
18. Plaseski T, Noveski P, Kocevska B, Dimitrovski C, Efremov GD, Plaseska-Karanfilska D. CAG repeat number in androgen receptor gene and male infertility. Balkan J Med Genet. 2007; 10(1): 19-24.
19. Plaseski T, Noveski P, Popeska Z, Efremov GD, Plaseska-Karanfilska D. Association study of single nucleotide polymorphisms in *FASLG, JMJDIA, LOC203413, TEX15, BRDT, OR2W3, INSR* and *TAS2R38* genes with male infertility. J Androl. 2011; 33(4): 675-683.
20. Tuttelmann F, Simoni M, Kliesch S, *et al.* Copy number variants in patients with severe oligozoospermia and Sertoli-cell-only syndrome. PLoS One. 2011; 6(4): e19426.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

HUMAN SEMINAL PLASMA PROTEOME STUDY: A SEARCH FOR MALE INFERTILITY BIOMARKERS

Davalieva K[1,]*, Kiprijanovska S[1], Noveski P[1], Plaseski T[2], Kocevska B[2], Plaseska-Karanfilska D[1]

***Corresponding Author:** Dr. Katarina Davalieva, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, 1000 Skopje, Republic of Macedonia; Tel.: +389-2-3235-410; Fax: +389-2-3115-434; E-mail: katarina@manu.edu.mk

ABSTRACT

Seminal plasma is a potential source of biomarkers for many disorders of the male reproductive system including male infertility. Knowledge of the peptide and protein components of seminal fluid is accumulating especially with the appearance of high-throughput MS-based techniques. Of special interest in the field of male infertility biomarkers, is the identification and characterization of differentially expressed proteins in seminal plasma of men with normal and impaired spermatogenesis. However, the data obtained until now is still quite heterogeneous and with small percentage of overlap between independent studies. Extensive comparative analysis of seminal plasma proteome is still needed in order to establish a potential link between seminal plasma proteins and male infertility.

Keywords: Biomarkers, Male infertility, Mass spectrometry (MS), Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF), Seminal plasma, Two-dimensional differential in-gel electrophoresis (2-D DIGE), Two- dimensional polyacrylamide gel electrophoresis (2-D PAGE).

[1] Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, 1000 Skopje, Republic of Macedonia

[2] Endocrinology and Metabolic Disorders Clinic, Faculty of Medicine, Skopje, Republic of Macedonia

INTRODUCTION

The proteomes most likely to contain clinically useful disease biomarkers are those of human body fluids. Proteomics has raised great expectations for the discovery of biomarkers for improved diagnosis or stratification of a wide range of diseases [1]. Blood plasma and other body fluids, which also include seminal fluids, are expected to be excellent sources of protein biomarkers because they circulate through, or come in contact with a variety of tissues, and during this contact they are likely to pick up proteins secreted or shed by tissues, a hypothesis that has been tested and confirmed [2].

In addition to the general physiological importance of knowing the composition of seminal fluid, medical interest centers on two main areas: infertility and prostate cancer. Male infertility is a widespread medical condition with large societal and emotional costs. Between one in six and one in 10 couples seek medical help for the problem of sub fertility. In 20.0-25.0% of cases the infertility problem is due to the male partner, in 30.0-40.0% the problem is predominantly female; in approximately 30.0% of cases, abnormalities are found in both partners, and in 15.0% no specific factor is identified [3]. Since seminal fluid has an important role in spermatozoa survival and overall fertilization success, its impairment can be directly connected to infertility.

Seminal plasma contains many distinct protein components that are important for the functioning and survival of spermatozoa. It is a mixture of secretions from several male accessory glands, including prostate, seminal vesicles, epididymis, and Cowper's gland. The average protein concentration of human seminal plasma ranges from 35 to 55 g/L making it a rich, as well as an easily accessible source for protein identification. Seminal plasma has the feature common to many other body fluids, namely it is characterized by a high dynamic range of protein abundance, making low-abundance components difficult to analyze.

Characterization of Human Seminal Plasma Protein and Peptide Constituents. Research of seminal plasma proteome began using two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) and the electrophoretic transfer of proteins to nitrocellulose sheets followed by subsequent detection with immunological procedures in 1981 [4]. By this approach, a few seminal plasma proteins such as prostatic acid phosphatase (PAP), some glycoproteins and creatinin kinases, were identified. Two-dimensional (2-D) gels coupled with mass spectrometry (MS) were applied to study the role of seminal plasma proteins in impaired spermatogenesis in 2001, and about 750 spots were detected in the 2-D maps of seminal plasma from a fertile men [5]. However, only two seminal plasma constituents, namely, PAP and prostatic specific antigen (PSA) were identified by matrix-assisted laser desorption/ ionization time of flight (MALDI-TOF) in this study, as the two major spot clusters in the 2-D map of the seminal plasma proteome. By using 2-D PAGE with MALDI and liquid chromatography-electrospray ionization (LC-ESI), Fung and colleagues [6] identified 100 different protein constituents in seminal plasma. This study established that there were multiple post-translational variants of the majority of the proteins. Overall, hormones, growth factors and bio-active peptides were detected and identified. Pilch and Mann [7] successfully catalogued 932 proteins in seminal plasma using Fourier transform MS and two consecutive stages of MS fragmentation. This study identified proteins known to be characteristic for each of the organs contributing to the formation of seminal plasma: prostate, seminal vesicles, epididymis, and the bulbourethral gland. The large proportions of the identified proteins were extracellular, secreted by the male accessory glands as well as extracellular matrix proteins that are required for the classical functions of seminal fluid. The second class of the identified proteins was found to originate from prostasomes, while the third class of proteins was the result of an epithelial shredding. The overall numbers and proportions of proteins in the identified proteome indicated that the predominant functions are in clot formation and liquefaction, in metabolic support and protection for the spermatozoa and in immunological reactions. The identification of seminal plasma protein and peptide constituents by using combined 2-D PAGE/ MS/MS and 2D nano high performance liquid chromatography (HPLC)/MS/MS is currently in progress at our institution.

Identification of some well defined subgroup of human seminal plasma proteins such as heparin-binding proteins that are included in the fertilization process were also carried out by the use of affinity chromatography, 2D electrophoresis and MALDI-TOF MS [8]. Difference in the expression of sperm membrane proteins between fertile and infertile males [9], as well as proteins associated with human fertilization capability [10], have also been investigated.

Comparative Analysis of Seminal Plasma Proteome With Identification of Potential Infertility Biomarkers. The research in the field of comparative proteomic analysis of male infertility-associated seminal plasma proteins has been carried out for almost a decade. However, extensive comparative analysis of seminal plasma proteome in order to establish a potential link between seminal plasma proteins and male infertility is still lacking.

By investigating the differential expression of proteins between two or more biologically different states of reduced male fertility or infertility, potential diagnostic or prognostic markers may be identified. The gold standard techniques for this is still 2-D PAGE/2-D DIGE coupled to MS because this approach is amenable to the separation and visualization of a wide range of proteins, together with their post translational variants.

Identification of 61 differentially expressed proteins based on 2-D PAGE and tandem MS analysis of seminal plasma of azoospermic and vasectomized patients has been reported by Sarita-Garibaldi and coworkers [11]. Two- dimensional DIGE combined with MS was also carried out to find seminal plasma

proteins associated specifically with azoospermia and revealed a total of four potential markers for non obstructive azoospermia [12]. The identified potential biomarkers were stabilin 2 (STAB2), 135 kd centrosomal protein (CP135), guanine nucleotide, releasing protein (GNRP), and prolactin-inducible protein (PIP). In a recent study, we utilized the 2-D DIGE/MS approach to detect differential protein expression of seminal plasma proteins between four distinct groups of men with normozoospermia, oligozoospermia, asthenozoospermia and azoospermia [13]. The results from our study demonstrated that normozoospermic, oligozoospermic and asthenozoospermic groups have similar seminal plasma protein profiles, resulting in no statistically significant differences in protein expression. However, in the azoospermic group there were a total of eight proteins with a statistically significant increase of expression in comparison with the rest of the studied groups. From the eight differentially expressed proteins, seven were successfully identified by MS as fibronectin (FINC), prostatic acid phosphatase (PAP), proteasome subunit α type-3 (PSA3), β-2-microglobulin (B2MG), galectin-3 binding protein (LG3BP), prolactin inducible protein (PIP) and cytosolic non specific dipeptidase (CNDP2). All of the reported proteins have already been reported as constituents of seminal plasma. Most of these proteins (FINC, PAP, B2MG, LG3BP, PIP) are localized in the extracellular region and have the molecular function of protein binding. From all the differentially expressed spots found in this study, PAP was found to be exclusively increased in azoospermic patients compared with the rest of the studied groups. Prostatic acid phosphatase is the most abundant phosphatase in human prostate tissue and has been studied extensively, primarily due to its clinical relevance as a biomarker of prostate carcinoma [14]. Several other studies have implicated PAP in male infertility before, with enzyme levels inversely correlated with sperm concentration [15-17].

Another study, based on the comparative analysis of the levels of PIP in fertile and infertile men, has proposed that PIP could be an immunoglobulin G-binding protein [18]. Prolactin inducible protein was shown to exist in several isoforms in seminal plasma by Western blot. The pattern of PIP isoform variability in seminal plasma from fertile and infertile men is quite complex but one multimeric form of PIP was confirmed to be absent from the seminal plasma of fertile men.

The levels of fibronectin (FINC fragments in seminal plasma samples were found higher in seminal plasma with abnormal semen characteristics than in the normozoospermic group [19]. The results suggest that seminal plasma FINC fragments may contribute to fertilization and the analysis of FINC fragmentation may have a diagnostic value in andrological investigations.

Proteomic analysis of seminal plasma from asthenozoospermia patients revealed a rich source of biomarker candidates for male infertility, proposing that functional abnormalities of the epididymis and prostate can contribute to asthenozoospermia [20]. Among the identified proteins, DJ-1, a protein that has been shown to be involved in the control of oxidative stress, was down regulated in asthenozoospermia patients and was proposed as a candidate biomarker for this condition.

In a conclusion, the research in the field of seminal plasma proteome and the search for biomarkers of male infertility is still ongoing. An in-depth understanding of the seminal plasma proteome would contribute greatly to the elucidation of the roles of seminal plasma proteins in the regulation of motility and to the establishment of biomarkers for male infertility. There have been different sets of proteins proposed as biomarkers in different conditions of reduced fertility and/or infertility. However, only small set of proteins such as PAP, PIP, FINC are found differentially expressed in male infertility by independent studies. One of the reasons for this very low percentage of overlap between independent studies is that different proteomics techniques and their combinations were used. Therefore, the data from these studies is quite heterogeneous. However, the appearance of high-throughput MS-based techniques allows more detailed investigation of the proteomes of interest, among which is human seminal plasma proteome, and holds promise on more reproducible results in the future.

ACKNOWLEDGMENTS

The proteomic study of seminal plasma performed was supported by grant CRP/MAC09-01 from the International Centre for Genetics Engineering and Biotechnology, Trieste, Italy (to

D. Plaseska-Karanfilska). We are grateful to Cedric Broussard for the mass spectrometry analysis that were performed at the Plate-forme Proteomique Paris 5 (3P5), Institut Cochin, Inserm U1016 UMR8104 and Université Paris Descartes Sorbonne Paris Cité, Paris, France. We are also grateful to Dr. Martina Marchetti-Deschmann for the training in Peptide Mass Fingerprinting that was performed during the stay of Dr. Katarina Davalieva at the Institute of Chemical Technologies and Analytics, Division: Instrumental Analytical Chemistry, Vienna University of Technology, Austria. Dr. Katarina Davalieva's 2-week stay was supported by the FP7 project No. 229458 from the European Commission.

REFERENCES

1. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann PH. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics. 2003; 4(4): 523-538.
2. Zhang H, Liu AY, Loriaux P, *et al.* Mass spectrometric detection of tissue proteins in plasma. Mol Cell Proteomics. 2007; 6(1): 64-71.
3. De Kretser DM, Baker HW. Infertility in men: recent advances and continuing controversies. J Clin Endocrinol Metab. 1999; 84(10): 3443-3450.
4. Edwards JJ, Tollaksen SL, Anderson NG. Proteins of human semen. I. Two-dimensional mapping of human seminal fluid. Clin Chem. 1981; 27(8): 1335-1340.
5. Starita-Geribaldi M, Poggioli S, Zucchini M, *et al.* Mapping of seminal plasma proteins by two-dimensional gel electrophoresis in men with normal and impaired spermatogenesis. Mol Hum Reprod. 2001; 7(8): 715-722.
6. Fung KY, Glode LM, Green S, Duncan MW. A comprehensive characterization of the peptide and protein constituents of human seminal fluid. Prostate. 2004; 61(2): 171-181.
7. Pilch B, Mann M. Large-scale and high-confidence proteomic analysis of human seminal plasma. Genome Biol. 2006; 7(5): R40.
8. Kumar V, Hassan MI, Tomar AK, *et al.* Proteomic analysis of heparin-binding proteins from human seminal plasma: a step towards identification of molecular markers of male fertility. J Biosci. 2009; 34(6): 899-908.
9. Rajeev SK, Reddy KVR. Sperm membrane protein profiles of fertile and infertile men: identification and characterization of fertility associated sperm antigen. Hum Reprod. 2004; 19(2): 234-242.
10. Zhang J, Wu J, Huo R, *et al.* ERp57 is a potential biomarker for human fertilization capability. Mol Hum Reprod. 2007; 13(9): 633-639.
11. Starita-Geribaldi M, Roux F, Garin J, Chevallier D, Fenichel P, Pointis G. Development of narrow immobilized pH gradients covering one pH unit for human seminal plasma proteomic analysis. Proteomics. 2003; 3(8): 1611-1619.
12. Yamakawa K, Yoshida K, Nishikawa H, Kato T, Iwamoto T. Comparative analysis of interindividual variations in the seminal plasma proteome of fertile men with identification of potential markers for azoospermia in infertile patients. J Androl. 2007; 28(6): 858-865.
13. Davalieva K, Kiprijanovska S, Noveski P, *et al.* Proteomic analysis of seminal plasma in men with different spermatogenic impairment. Andrologia. 2012; 44(4): 256-264.
14. Azumi N, Traweek ST, Battifora H. Prostatic acid phosphatase in carcinoid tumors. Immunohistochemical and immunoblot studies. Am J Surg Pathol. 1991; 15(8): 785-790.
15. Dave BN, Rindani TH. Acid phosphatase activity in human semen. Int J Fertil. 1988; 33(1): 45-47.
16. Vaubourdolle M, Clavel JP, Gonzales J, Galli A. Evaluation of acid phosphatase isoenzymes in seminal fluid from normozoospermic, oligozoospermic, azoospermic and asthenoteratozoospermic men. Andrologia. 1985; 17(6): 598-604.
17. Ziyyat A, Barraud-Lange V, Sifer C, Ducot B, Wolf JP, Soufir JC. Paradoxical increase of sperm motility and seminal carnitine associated with moderate leukocytospermia in infertile patients. Fertil Steril. 2008; 90(6): 2257-2263.
18. Chiu WW, Chamley LW. Human seminal plasma prolactin-inducible protein is an immunoglobulin G-binding protein. J Reprod Immunol. 2003; 60(2): 97-111.
19. Katnik-Prastowska I, Przybysz M, Chelmonska-Soyta A. Fibronectin fragments in human seminal plasma. Acta Biochim Pol. 2005; 52(2): 557-560.
20. Wang J, Zhang HR, Shi HJ, *et al.* Proteomic analysis of seminal plasma from asthenozoospermia patients reveals proteins that affect oxidative stress responses and semen quality. Asian J Androl. 2009; 11(4): 484-491.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

RAPID AND NON INVASIVE PRENATAL DIAGNOSIS

Madjunkova S[1], Sukarova-Stefanovska E[1], Kocheva S[2],Maleva I[1], Noveski P[1], Kiprijanovska S[1], Stankova K[3],Dimcev P[4], Madjunkov M[5], Plaseska-Karanfilska D[1,*]

***Corresponding Author**: Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)23235-410; Fax: +389 (0)2-3115-434; E-mail: dijana@manu.edu.mk

INTRODUCTION

Prenatal screening and diagnosis are routinely offered at antenatal care clinic visits, and are important in decision making about the continuation of pregnancies affected by genetic conditions for which there are no cures, and prevention through therapeutic abortion is a reasonable option. Prenatal screening is offered to all pregnant women and include fetal ultrasonography and maternal serum biochemistry to select the pregnancies at-risk for chromosomal abnormalities. However, these methods have limited sensitivities (60.0-75.0%) and specificities (false positive rate of 5.0%). Even when used in combination and taking into account maternal age, the identification rate of affected fetuses does not exceed 90.0% [one]. Prenatal diagnosis is usually preformed for detection of chromosomal aneuploidies or monogenic diseases in "high risk" pregnancies. Diagnostic testing currently requires a sample of fetal cells obtained either by chorionic villus sampling (CVS) between 10 and 14 weeks gestation or by amniocentesis after 15 weeks of gestation. However, these invasive procedures carry a risk of miscarriage of around 1.0% [2].

Prenatal Diagnosis of Chromosomal Abnormalities. Chromosomal abnormalities (numerical or structural) occur in 1 of 160 live births, with extra copies of chromosomes 21, 18, and 13 accounting for the majority of numerical alterations that are not related to sex chromosomes. The prevalence of trisomies is highest in the first trimester because of subsequent miscarriage and demise of aneuploid conceptuses during pregnancy [3]. Conventional cytogenetic techniques (karyotyping) are usually used to detect aneuploidies and large (5-10 Mb) rearrangements in fetal cells (amniocytes, trophoblasts), however, these are time-consuming (2-3 weeks), subjective (small rearrangements) and expensive. The development of molecular methods for the rapid, targeted detection of aneuploidies of chromosomes 13, 18, 21 and the sex chromosomes by quantitative fluorescent polymerase chain reaction (QF-PCR) [4,5] using fetal DNA, do not provide a genome-wide screen for unexpected imbalances, but are rapid (24-48 hours), accurate and inexpensive. Multiplex ligation probe amplification (MLPA) is a recent technique for relative quantita-

[1] Macedonian Academy of Sciences and Arts, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Skopje, Republic of Macedonia
[2] Pediatric Clinic, Medical Faculty, University "Stt. Cyril and Methodius", Skopje, Republic of Macedonia
[3] Private Health Organisation "Dr. Ina," Skopje, Republic of Macedonia
[4] Private Health Organisation "Sistina," Skopje, Republic of Macedonia
[5] Clinic for Gynaecology and Obstetrics, Faculty of Medicine, Skopje, Republic of Macedonia

tion of up to 40 to 45 nucleic acid targets. Several MLPA commercial kits are used for prenatal detection of common aneuploidies (chromosomes 13, 18, 21, X and Y), common microdeletion syndromes and subtelomeric copy-number changes, identification of marker chromosomes, and detection of familial copy-number changes in single genes [6-8]. The most powerful technique for genome wide screening is array comparative genomic hybridization (aCGH), which has the potential to combine the speed of DNA analysis with a large capacity to scan for subtle genomic abnormalities (approximately additional 10.0% of karyotying) respective to the resolution of the used arrays [9-11], but is expensive, time-consuming and requires a high degree of expertise.

Non Invasive Prenatal Diagnosis. The discovery of cell-free fetal DNA (cffDNA) in maternal plasma in 1997 opened up new avenues for prenatal diagnosis [12,13]. Fractional concentration of fetal DNA is ~10.0%, coexists with a background of maternal DNA and is present in maternal plasma from approximately the 6th gestational week [14]. Techniques, such as real-time PCR (ReTi-PCR) and digital PCR, provide sufficient sensitivity for reliable non invasive assessment of this cffDNA pool for paternally inherited traits such as sex and RHD status, offering possibilities for non invasive prenatal diagnosis of X-linked disorders (such as Duchenne/Becker muscular dystrophy, Hemophilia A, Hemophilia B, *etc.*) and RhD incompatibility, respectively [15,16]. By detecting the presence of fetal-specific paternally inherited mutant alleles in maternal plasma, diagnosis of autosomal dominant diseases transmitted by the father could be made non invasively, whereas the absence of such alleles could be used to exclude fetal inheritance of autosomal recessive diseases [14,17-20]. Quantification of cffDNA, specific fetal and maternal DNA and mRNA single nucleotide polymorphism allelic ratios have been used to detect fetal aneuploidies, however, the limitations of these techniques affect the accuracy of the diagnosis [21-23]. Improvements were made after the discovery of the unmethylated *SEPINB5* gene that turned out to be the first sex- and polymorphism-independent fetal DNA marker found in maternal plasma [24-27]. The differential methylation of placenta and maternal blood provides a rich source of markers for non invasive prenatal diagnosis, however, further research is needed to render the techniques widely applicable. Implementing the new and robust next generation sequencing techniques in detection of the fetal aneuploidy made the detection for Down's syndrome to have 98.6-100.0% sensitivity and 96.8-97.9% specificity [28,29].

Prenatal Diagnosis of Monogenic Diseases. Monogenic diseases are the second most frequent indication for prenatal diagnosis. The incidence of these diseases, depending on the population, is up to 2.0% newborns. Although there are some biochemical tests and ultrasound findings to screen and identify pregnancies at-risk for specific monogenetic disorders, still the diagnosis is usually established after the fetus is born in couples with no familial history of the disease. In families at-risk for monogenic disease, prenatal diagnosis is used to determine fetal health and to provide adequate management of the pregnancy and prenatal or perinatal treatment. The new developments in prenatal testing using cffDNA and their translation into clinical practice are going to make a difference in selection of pregnancies at-risk for monogenic disorders that need invasive testing.

Prenatal Diagnosis at the Research Centre for Genetic Engineering and Biotechnology (RCGEB)"Georgi D. Efremov," Skopje, Republic of Macedonia. In the last 20 years, the researchers at the RCGEB "Georgi D. Efremov" have performed more than 80 prenatal diagnoses for different monogenic diseases, such as hemoglobinopathies, cystic fibrosis, Duchenne/Becker muscular dystrophy, spinal muscular atrophy, hemophilia A, Lesch Nyhan syndrome, Rett syndrome, phenylketonuria, galactosemia, pseudohypoaldosteronism, *etc.* [30,31]. The prenatal diagnosis was performed on fetal DNA by using standard molecular genetic techniques for direct diagnosis of the disease or by using informative polymorphic DNA markers for indirect diagnosis.

In 2001, the rapid prenatal detection of the most common chromosomal aneuploidies (chromosomes 13, 18, 21, X and Y) by the multiplex QF-PCR (mQF-PCR) method was introduced at the RCGEB "Georgi D. Efremov" [32,33]. We have developed a one-tube mQF-PCR assay for amplification of 22 highly polymorphic short tandem repeat (STR) markers (at least four by analyzed chromosome) (Table 1). Since then, more than 2200 prenatal diagnoses of common aneuploidies in at-risk preg-

Table 1. Short tandem repeat markers used in one-tube quantitative fluorescent polymerase chain reaction assay for detection of aneuploidies of chromosomes 13, 18, 21, X and Y.

	Chromosome 13	Chromosome 18	Chromosome 21	Chromosome X	Chromosome Y
STR markers	D13S631	D18S535	D21S1414	AMEL X/Y[a]	AMEL X/Y[a]
	D13S305	D18S391	D21S1446	DXYS218[a]	DXYS218[a]
	D13S258	D18S386	D21S1411	TAF9B[b]	MYPT2Y[c]
	D13S1817	D18S390	D21S1435	XHPRT	SRY
				DXS6803	DYS448

[a] Locus on chromosomes X and Y.
[b] Locus on chromosomes X and 3.
[c] Locus on chromosomes Y and 2.

nancies have been performed using the mQF-PCR assay as a stand-alone test [34]. It was also used in the prenatal cases of monogenic diseases to control maternal contamination of the fetal material. The prenatal diagnosis was performed on genomic DNA isolated from fetal cells collected by amniocentesis or CVS. Maternal blood samples were analyzed in all blood contaminated amniotic samples and in most chorionic villi samples. No discordant results were obtained when cytogenetic analysis was performed in addition to QF-PCR. Polymorphic duplications involving STR markers D13S631, D21S1441, D18S978 or D18S535 were detected in seven fetuses; in all fetuses the duplications were inherited from one of the parents. Using this method we were also able to determine the parental origin of the aneuploidy [35,36]. In our experience, the QF-PCR method is an efficient, rapid and reliable method for prenatal diagnosis of the most common chromosome aneuploidies. In addition, it can provide information about the origin of the aneuploidy and maternal contamination of the fetal material.

In some "high risk" pregnancies with normal QF-PCR results, we have used MLPA kits to analyze subtelomeric regions and common microdeletion syndromes. In addition to this, aCGH has been employed in prenatal diagnosis of a few fetuses with specific abnormal ultrasound findings.

We have also evaluated the specificity and sensitivity of the real-time quantitative PCR method for non invasive fetal sex determination using cffDNA from maternal plasma. Our initial results showed that this is a promising approach for fetal gender determination in pregnancies at-risk for a fetus with an X-linked disorder [37]. Our recent study of non invasive determination of fetal RHD status, using cffDNA from maternal plasma in RhD negative pregnant women, showed 100.0% concordant results with those obtained on fetal DNA from amniocytes or CVS. This is a promising test that can be used in clinical practice for targeted anti-RhD prophylaxis and improvement of management of RHD fetomaternal incompatibility. Using a multi copy marker on Y chromosome (DYS14), we have increased the sensitivity and specificity of the non invasive fetal sex determination using cffDNA. This method will be used in the future for non invasive fetal sex determination in pregnancies at-risk for X-linked disorders. Our further plans include translation of the non invasive tests using cffDNA for diagnosis of monogenic disorders and chromosomal aneuploidies into clinical practice.

ACKNOWLEDGMENTS

We are grateful to Professor Neil D. Avent for the study including non invasive prenatal determination of fetal sex and RHD status using cffDNA from maternal plasma from pregnant Macedonian women, that was performed during the stay of Dr. Svetlana Madjunkova at the School of Biological and Biomedical Sciences, University of Plymouth, Plymouth, Devon, UK. Dr. Madjunkova's 2-week stay was supported by the FP7 project No. 229458 from the European Commission.

REFERENCES

1. Kagan KO, Wright D, Spencer K, Molina FS, Nicolaides KH. First-trimester screening for trisomy 21 by free β-human chorionic gonadotropin and pregnancy-associated plasma protein-A: impact of mater-

nal and pregnancy characteristics. Ultrasound Obstet Gynecol. 2008; 31(5): 493-502.
2. Mujezinovic F, Alfirevic Z. Procedure-related complications of amniocentesis and chorionic villus sampling: a systematic review. Obstet Gynecol. 2007; 110(3): 687-694.
3. Driscoll DA, Gross S. Clinical practice. Prenatal screening for aneuploidy. N Engl J Med. 2009; 360(24): 2556-2562.
4. Cirigliano V, Ejarque M, Canadas MP, *et al.* Clinical application of multiplex quantitative fluorescent polymerase chain reaction (QF-PCR) for the rapid prenatal detection of common chromosome aneuploidies. Mol Hum Reprod. 2001; 7(10): 1001-1006.
5. Faas BH, Cirigliano V, Bui TH. Rapid methods for targeted prenatal diagnosis of common chromosome aneuploidies. Semin Fetal Neonatal Med. 2011; 16(2): 81-87.
6. Willis AS, van den Veyver I, Eng CM. Multiplex ligation-dependent probe amplification (MLPA) and prenatal diagnosis. Prenat Diagn. 2012; 32(4): 315-320.
7. Bocian E, Kasprzycka J, Jakubow-Durska K, Luszczek A, Bernaciak J. [Usefulness of MLPA technique for rapid prenatal detection of aneuploidy. Results of 409 diagnostic studies]. Ginekol Pol. 2011; 82(9): 680-684.
8. Konialis C, Hagnefelt B, Sevastidou S, *et al.* Uncovering recurrent microdeletion syndromes and subtelomeric deletions/duplications through non-selective application of a MLPA-based extended prenatal panel in routine prenatal diagnosis. Prenat Diagn. 2011; 31(6): 571-577.
9. Rickman L, Fiegler H, Shaw-Smith C, *et al.* Prenatal detection of unbalanced chromosomal rearrangements by array CGH. J Med Genet. 2006; 43(4): 353-361.
10. Brady PD, Devriendt K, Deprest J, Vermeesch JR. Array-based approaches in prenatal diagnosis. Methods Mol Biol. 2012; 838: 151-171.
11. Fiorentino F, Caiazzo F, Napolitano S, *et al.* Introducing array comparative genomic hybridization into routine prenatal diagnosis practice: a prospective study on over 1000 consecutive clinical cases. Prenat Diagn. 2011; 31(13): 1270-1282.
12. Dennis Lo YM, Chiu RW. Prenatal diagnosis: progress through plasma nucleic acids. Nat Rev Genet. 2007; 8(1): 71-77.
13. Lo YM, Corbetta N, Chamberlain PF, *et al.* Presence of fetal DNA in maternal plasma and serum. Lancet. 1997; 350(9076): 485-487.
14. Lun FM, Tsui NB, Chan KC, *et al.* Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. Proc Natl Acad Sci USA. 2008; 105(50): 19920-19925.
15. Hromadnikova I, Vechetova L, Vesela K, Benesova B, Doucha J, Vlk R. Non-invasive fetal RHD and RHCE genotyping using real-time PCR testing of maternal plasma in RhD-negative pregnancies. J Histochem Cytochem. 2005; 53(3): 301-305.
16. Tsui NB, Kadir RA, Chan KC, *et al.* Noninvasive prenatal diagnosis of hemophilia by microfluidics digital PCR analysis of maternal plasma DNA. Blood. 2011; 117(13): 3684-3691.
17. Saito H, Sekizawa A, Morimoto T, Suzuki M, Yanaihara T. Prenatal DNA diagnosis of a single-gene disorder from maternal plasma. Lancet. 2000; 356(9236): 1170.
18. Chiu RW, Lau TK, Leung TN, Chow KC, Chui DH, Lo YM. Prenatal exclusion of β thalassaemia major by examination of maternal plasma. Lancet. 2002; 360(9338): 998-1000.
19. Li Y, Di Naro E, Vitucci A, Zimmermann B, Holzgreve W, Hahn S. Detection of paternally inherited fetal point mutations for β-thalassemia using size-fractionated cell-free DNA in maternal plasma. JAMA. 2005; 293(7): 843-849.
20. Legler TJ, Muller SP, Haverkamp A, Grill S, Hahn S. Prenatal RhD testing: a review of studies published from 2006 to 2008. Transfus Med Hemother. 2009; 36(3): 189-198.
21. Zhong XY, Burk MR, Troeger C, Kang A, Holzgreve W, Hahn S. Fluctuation of maternal and fetal free extracellular circulatory DNA in maternal plasma. Obstet Gynecol. 2000; 96(6): 991-996.
22. Chiu RW, Chan KC, Gao Y, *et al.* Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc Natl Acad Sci USA. 2008; 105(51): 20458-20463.
23. Tsui NB, Akolekar R, Chiu RW, *et al.* Synergy of total PLAC4 RNA concentration and measurement of the RNA single-nucleotide polymorphism allelic ratio for the noninvasive prenatal detection of trisomy 21. Clin Chem. 2010; 56(1): 73-81.
24. Tong YK, Chiu RW, Akolekar R, *et al.* Epigenetic-genetic chromosome dosage approach for fetal trisomy 21 detection using an autosomal genetic reference marker. PLoS One. 2010; 5(12): e15244.
25. Chiu RW, Chim SS, Wong IH, *et al.* Hypermethylation of RASSF1A in human and rhesus placentas. Am J Pathol. 2007; 170(3): 941-950.
26. Chim SS, Jin S, Lee TY, *et al.* Systematic search for placental DNA-methylation markers on chromosome 21: toward a maternal plasma-based epigenetic test for fetal trisomy 21. Clin Chem. 2008; 54(3): 500-511.

27. Brown L, Brown G, Vacek P, Brown S. Aneuploidy detection in mixed DNA samples by methylation-sensitive amplification and microarray analysis. Clin Chem. 2010; 56(5): 805-813.
28. Palomaki GE, Kloza EM, Lambert-Messerlian GM, *et al.* DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. Genet Med. 2011; 13(11): 913-920.
29. Chiu RW, Akolekar R, Zheng YW, *et al.* Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. BMJ. 2011; 342:c7401.
30. Kocheva S, Trivodalieva S, Plaseska-Karanfilska D, Vlaski-Jekic S, Kuturec M, Efremov GD. Prenatal diagnosis in Macedonian Duchenne muscular distrophy families. Balkan J Med Genet. 2008; 11(2): 59-63.
31. Kocheva SA, Plaseska-Karanfilska D, Trivodalieva S, Kuturec M, Vlaski-Jekic S, Efremov GD. Prenatal diagnosis of spinal muscular atrophy in Macedonian families. Genet Test. 2008; 12(3): 391-393.
32. Arsovska S, Dimcev P, Kaeva M, Efremov GD, Plaseska-Karanfilska D. Rapid prenatal diagnosis of common aneuploidies by quantitative fluorescent polymerase chain reaction. Proceedings of the 5th Balkan Meeting on Human Genetics, Sofia, Bulgaria, August 28-September 1, 2002; 145.
33. Plaseska-Karanfilska D, Talaganova S, Trivodalieva S, Efremov GD. Prenatal diagnosis of aneuploidies of chromosomes 13, 18, 21, X and Y by QF-PCR in the Republic of Macedonia. Proceedings of the European Human Genetics Conference, Nice, France, June 16-19, 2007. Eur J Hum Genet. 2007; 140.
34. Plaseska-Karanfilska D, Madjunkova S, Maleva I, Kiprijanovska S. Rapid prenatal diagnosis of common chromosome aneuploidies using quantitative fluorescent (QF)-PCR: 10 years experience in a center from the Republic of Macedonia. Proceedings of the 12th International Congress of Human Genetics, Montreal, Canada, October 11-15, 2011; 1369W.
35. Talaganova S, Trivodalieva S, Arsovska S, Kaeva M, Efremov GD, Plaseska-Karanfilska D. The parental and meiotic origin of extra chromosome 21 in Down syndrome patients. Proceedings of the 7th Balkan Meeting on Human Genetics, Skopje, Macedonia, August 31-September 2, 2006. Balkan J Med Genet. 2006; 118.
36. Kiprijanovska S, Trivodalieva S, Gefremov GD, Plaseska-Karanfilska D. The parental origin of chromosome aneuploidies. Proceedings of the 8th Balkan Meeting of Human Genetics, Cavtat, Croatia, May 14-17, 2009; 23.
37. Davalieva K, Dimcev P, Efremov GD, Plaseska-Karanfilska D. Non-invasive fetal sex determination using real-time PCR. J Matern Fetal Neonatal Med. 2006; 19(6): 337-342.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

INTEGRATIVE 'OMIC' APPROACH TOWARDS UNDERSTANDING THE NATURE OF HUMAN DISEASES

Peterlin B*, Maver A

***Corresponding Author:** Professor Borut Peterlin, M.D., Ph.D., Institute of Medical Genetics, Department of Gynecology and Obstetrics, University Medical Centre Ljubljana, 3, Šlajmerjeva Street, Ljubljana 1000, Slovenia; Tel./Fax: +386(0)1-540-1137; E mail: borut.peterlin@guest.arnes.si

ABSTRACT

The combination of improving technologies for molecular interrogation of global molecular alterations in human diseases along with increases in computational capacity, have enabled unprecedented insight into disease etiology, pathogenesis and have enabled new possibilities for biomarker development. A large body of data has accumulated over recent years, with a most prominent increase in information originating from genomic, transcriptomic and proteomic profiling levels. However, the complexity of the data made discovery of high-order disease mechanisms involving various biological layers, difficult, and therefore required new approaches toward integration of such data into a complete representation of molecular events occurring on cellular level.

For this reason, we developed a new mode of integration of results coming from heterogeneous origins, using rank statistics of results from each profiling level. Due to the increased use of next-generation sequencing technology, experimental information is becoming increasingly more associated to sequence information, for which reason we have decided to synthesize the heterogeneous results using the information of their genomic position. We therefore propose a novel positional integratomic approach toward studying 'omic' information in human disease.

Keywords: Data integration, Genomics, Transcriptomics, High-throughput technologies

INTRODUCTION

The development of microarray technology in the last decade and the upsurge of next-generation sequencing in the last few years has provided an abundance of data originating from various biological levels, most prominently from genomic and transcriptomic levels [1,2]. Such novel approaches have contributed greatly towards our understanding of physiological cellular processes, as well as molecular changes that occur in human disease. The high-dimensional nature of data originating from these studies has also opened an array of new theoretical and statistical challenges that had to be faced in order to attain acceptable reproducibility and consistency of scientific results [3]. In particular, a large number of hypotheses tested in a single experiment produced a substantial amount of statistical noise, causing large numbers of false-positive detections and undue omission of many true-positive results. Although statistical methods have been developed to address these issues, difficulties in in-

Clinical Institute of Medical Genetics, Department of Obstetrics and Gynecology, University Medical Centre Ljubljana, 3 Šlajmerjeva Street, Ljubljana 1000, Slovenia

creasing specificity and sensitivity of highly parallel approaches still exist, with the greatest notoriety in the field of human diseases belonging to a group of common, complex disorders.

In an attempt to alleviate these drawbacks, we developed a method that harnesses the biological relations between data originating from studies investigating human disease on various biological levels. An example of such an approach may be illustrated by the fact that genomic alterations associated with human disease, *i.e.,* multiple sclerosis (MS), are usually investigated and interpreted separately from transcriptomic alterations occurring in MS. The biological relation between these two layers may thus be utilized to favor prioritization of genes that were detected on both layers, therefore reducing noisy results and facilitating detection of true biological data. We expect that with the inclusion of increasing the number of biological layers and increasing the number of studies in the database used for integration, the comprehensiveness and biological validity of prioritized genes would increase progressively.

MATERIALS AND METHODS

The pathway towards constructing the initial database used for subsequent integration is highly dependent on the disease of interest. While some common disorders have been investigated in several 'omic' studies that investigated several biological cellular levels, the sourcing data for other diseases may be more scarce. The search for data sources should be initiated by an overview of literature published to date. When the investigator is familiar with the studies performed, the published reports and their tables in supplemental materials may be used to extract the lists of genes or other genomic features with detected significant alterations.

A crucial step in obtaining data sources of high quality is inspection of available databases that are stored in public data repositories. These tend to be highly specialized for the biological layer of investigation. For genomic data from genome-wide association studies, data may be extracted from dbGAP (http://www.ncbi.nlm.nih.gov/gap), for epigenomic, transcriptomic and methylomic data, Array Express (http://www.ebi.ac.uk/arrayexpress/) or Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/ geo/), and for next generation sequencing databases European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ ena/) and Sequence Read Archive (SRA; http://www.ncbi. nlm.nih.gov/ sra) [4-7].

After all the sources have been investigated, a collected database of features [genes, mRNAs, microRNAs (miRNAs), CpG islands, proteins and others] with significant alterations in chosen disease should be prepared for each included study. We also advise collection of information, such as significance values and fold change values, on which prioritization of features for each biological layer will be performed in the later steps. If the latter information is not available, all the significant alterations in a given study will have the same importance in integration. In the following section, significant results from various study types will be collectively referred to as "signals" for reasons of clarity.

Data Integration. Before data can be integrated, they have to be reduced to a universal common denominator. Due to increasing heterogeneity of genetic information, tying biological information to gene-level annotation is becoming increasingly more difficult. Genomic variation and methylation patterns are two examples of information that is prohibitively difficult to associate with genes in any straightforward manner, as such alterations occur in genes, between genes or spread across several genes. For this reason, we opted for an integration based on the genomic position of features originating from various data sources. This required the signals from all databases to be converted to their genomic positions and projected on the genome assembly backbone. This step then allows for complete omission of difficult annotation conversion steps, required before final integration can be performed, greatly simplifying the synthesis of heterogeneous data.

After signals are positioned on the genomic backbone, the complete assembly is divided into bins of equal size. For each study, a score is given to each of the bins, depending on the score of alterations residing in that segment of the genome. After this step, the scores of all bins are prioritized and their rank scores calculated. The integration step is attained when the non parametric rank product for each of the bins is calculated, and on the basis of rank scores of bins originating from each data source, as we have previously described [8]. The lower final rank product signifies that higher ranks

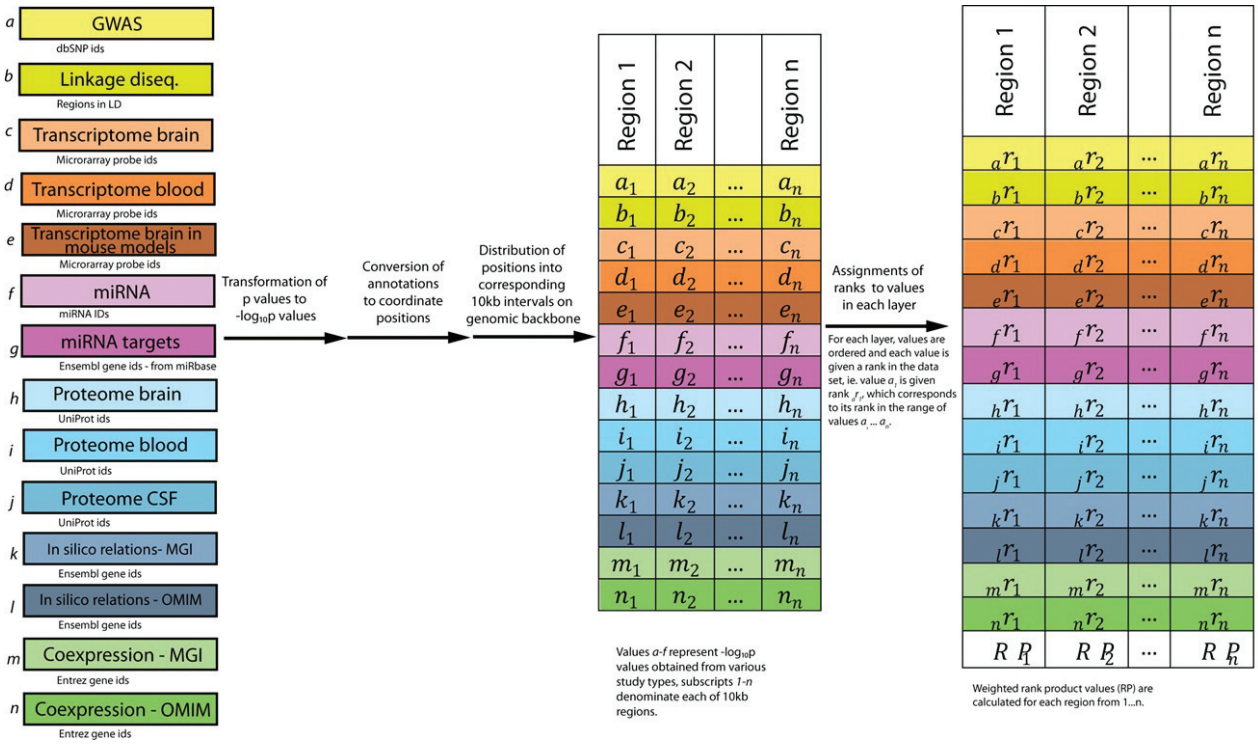


Figure 1. Process of integration of numerous heterogeneous data sources. First, data on significant alterations on a certain biological layer is obtained from selected studies (data from various layers is coded by letters a-n and differing colors). These alterations or signals are then positioned into genomic bins of fixed size and bin-scores for each of the bins is estimated. For each of the layers in a-n, bins are then prioritized on the basis of this score and the rank of each bin is separated. The final integration step is then performed by calculating rank products for each of the genomic bins, based on their rank in each of data sources.

were attained by bins on several separate biological layers [9]. Therefore, these bins represent genomic regions where accumulation of signals is detected on various biological levels, and thus represent regions of interest for further investigation. Ultimately, a permutational test may be employed to determine the significance of signal accumulation in each bin [8]. The detailed overview of the process may be observed in Figure 1.

RESULTS AND DISCUSSION

Results originating from the positional integratomic approach represent a prioritized list of genomic regions, where regions containing the greatest accumulation of heterogeneous biological alterations in an investigated disease rank highest and are characterized by lowest permutation test *p* values. As the integrative approach is performed for regions (bins) across the whole genome, the resulting genome-wide distribution of results from integration of data in human disease may be inspected. Genome-wide distribution of integration results for MS as an example of a complex autoimmune human disorder is represented in Figure 2. Here, the greatest accumulation of signals is observed on chromosome 6, specifically in the well-known human leukocyte antigen (HLA) region, suggesting that data from heterogeneous biological sources of 'omic' data indicate the role of this region in MS. Moreover, other regions have also attained high integration scores, suggesting importance of non-HLA regions in MS. Specifically, a region containing an interleukin-7 receptor gene (IL7R) attained very high integrative scores, not only on the basis of detections from genome-wide association studies, but also on the basis of evidence from expression profiling studies in blood and brain tissues. Additionally, the same region has been ranked high due to information obtained from various bioinformatic sources of data, such as KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways and co-expression information [10,11]. Such a heterogeneous body of evidence offers information of great relevance to true biological disease alterations and thus provides plausible candidate selection for further studies.

Figure 2. The genome-wide distribution of significance values, based on the permutation test of integration scores. Each region or genomic bin is represented by a dot whose height represent significances in the $-\log_{10}P$ form, with regions characterized by high accumulation of heterogeneous data attaining higher $-\log_{10}P$ values. The HLA region on chromosome 6 attained the highest score in these analyses with *p* values below $1 \bullet 10^{-9}$. Notably, non-HLA regions score high as well, offering a landscape of new genomic regions for further down-stream investigations

The positional approach offers great flexibility and control over parameters on which the final prioritization of genomic regions is based. Based on scientific questions, a researcher may be more interested in a contribution of only selected biological layers to the final integration score. For this reason, we have implemented means to allow custom weighting of different sources of data. For example, if one is interested in the relation between genomic variation and differential methylation, one may attribute those two sources greater weights and regions where signals from GWAS (genome-wide association studies), and global methylation studies aggregate will be obtained. Additional levels of control may also be obtained by customizing the size of genomic bins, allowing for detection of interactions that spread across larger genomic regions.

There has been great interest in deciphering the genetic factors with medium-to-low effect sizes as the explanation for the phenomenon of missing heritability in MS and other complex disorders [12,13]. Here, an integrative approach may assist in promoting detection of the genomic variant with its actual role in such complex disorder, and distinguishing them from spurious noise originating from statistical noise generated in genome-wide association studies. As large-scale studies, which attempt to detect low-effect susceptibility factors in human disease, have to be performed on large sample sizes, requiring great resources and effort [14], this approach may be a mode of comprehensive evidence-based selection of molecular determinants to investigate in such downstream validation studies.

With continuing development of high-throughput technologies, it is expected that the amount of the resulting data in large databases will continue to rise. For this reason, novel approaches for interpretation and understanding will also have to be prepared to face these challenges. As it is difficult for a single researcher or research group to have a comprehensive overview over such a vast information landscape, new means of presentation and access to these results will have to be envisaged. A position-based, integrative approach not only represents the means to quick insight into heterogeneous evidence from several large-scale studies, but is also a basis

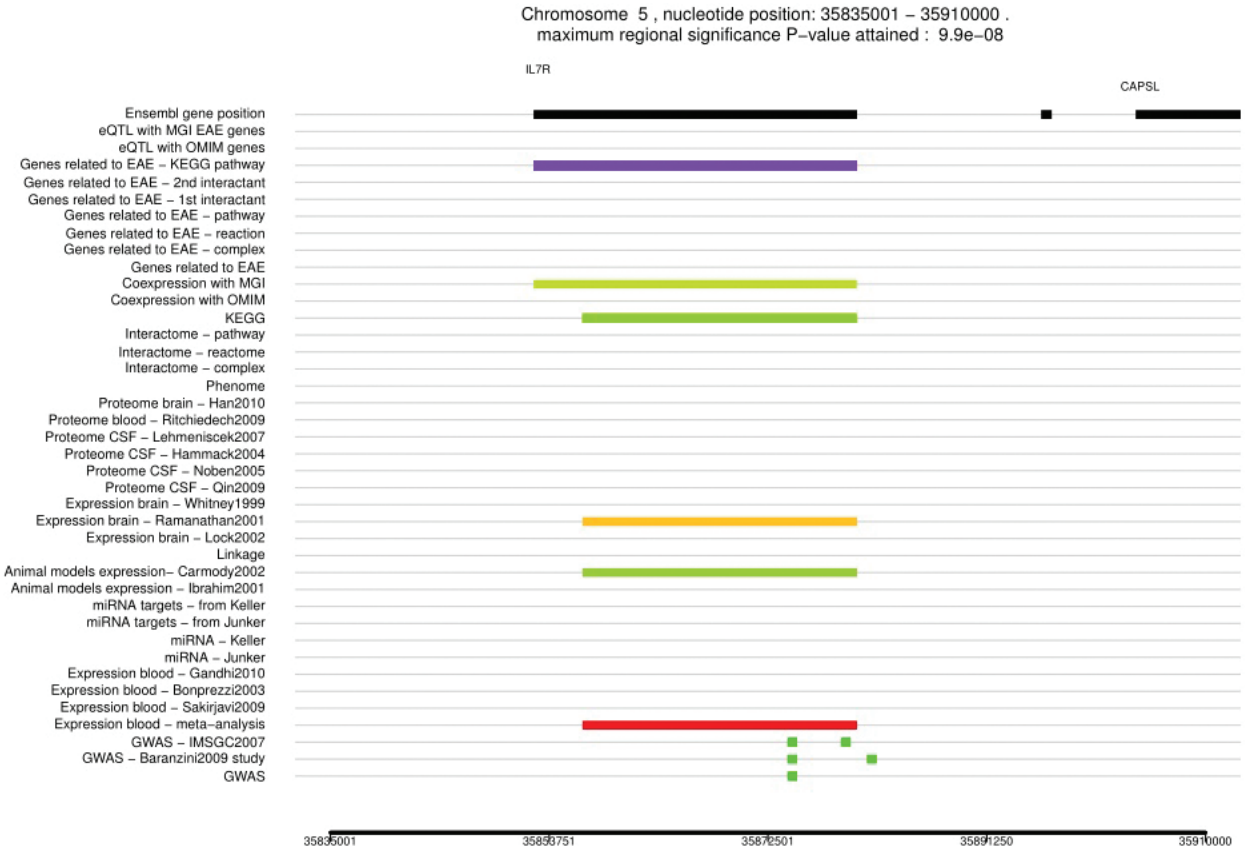


Figure 3. The supporting evidence substantiating the high score of the IL7R region on chromosome 5. Although the genetic variants in the IL7R gene have been detected in genome-wide association scans, there is also a substantial body of data supporting its relevance in MS. This support originates from whole-genome expression profiling studies in blood (red color) and brain (yellow color) as well as in expression profiling of brain samples from experimental autoimmune encephalitis animal models (dark green color). Additional in-silico relations, such as KEGG pathway relations and co-expression data support its relevance in MS.

toward the preparation of an interactive genome browser-like solutions for fast and easy access to this body of information.

REFERENCES

1. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet. 2006; 7(1): 55-65.
2. Schuster SC. Next-generation sequencing transforms today's biology. Nat Methods. 2008; 5(1): 16-18.
3. Shi L, Reid LH, Jones WD, *et al.* The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol. 2006; 24(9): 1151-1161.
4. Leinonen R, Sugawara H, Shumway M. The sequence read archive. Nucleic Acids Res. 2011; 39(Database issue): D19-D21.
5. Leinonen R, Akhtar R, Birney E, *et al.* The European nucleotide archive. Nucleic Acids Res. 2011; 39 (Database issue): D28-D31.
6. Parkinson H, Kapushesky M, Shojatalab M, *et al.* ArrayExpress–a public database of microarray experiments and gene expression profiles. Nucleic Acids Res. 2007; 35(Database issue): D747-D750.
7. Barrett T, Edgar R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. Methods Enzymol. 2006; 411: 352-369.
8. Maver A, Peterlin B. Positional integratomic approach in identification of genomic candidate regions for Parkinson's disease. Bioinformatics. 2011; 27(14): 1971-1978.
9. Breitling R, Herzyk P. Rank-based methods as a non-parametric alternative of the T-statistic for the analysis of biological microarray data. J Bioinform Comput Biol. 2005; 3(5): 1171-1189.
10. Kotera M, Hirakawa M, Tokimatsu T, Goto S, Kanehisa M. The KEGG databases and tools facilitating omics analysis: latest developments involving human diseases and pharmaceuticals. Methods Mol Biol. 2012; 802: 19-39.

11. Obayashi T, Kinoshita K. COXPRESdb: a database to compare gene coexpression in seven model animals. Nucleic Acids Res. 2011; 39(Database issue): D1016-D1022.
12. Sawcer S, Hellenthal G, Pirinen M, *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011; 476(7359?): 214-219.
13. Manolio TA, Collins FS, Cox NJ, *et al.* Finding the missing heritability of complex diseases. Nature. 2009; 461(7265): 747-753.
14. Marchini J, Donnelly P, Cardon LR. Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet. 2005; 37(4): 413-417.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

A PRELIMINARY microRNA ANALYSIS OF NON SYNDROMIC THORACIC AORTIC ANEURYSMS

Patuzzo C[1,*], Pasquali A[1], Malerba G[1], Trabetti E[1],
Pignatti PF[1], Tessari M[2], Faggian G[2]

***Corresponding Author:** Dr. Cristina Patuzzo, Department of Life and Reproduction Sciences, University of Verona, Strada Le Grazie 8, 37134,Verona, Italy; Tel.: +39-45-802-7207; Fax: +39-45-802-7180; E-mail: cristina.patuzzo@univr.it

ABSTRACT

The development of thoracic aortic aneurysms (TAAs) involves a multifactorial process resulting in alterations of the structure and composition of the extracellular matrix (ECM). Recently, modifications in microRNA (miRNA) expression were implicated in the pathogenesis of TAA.

This study presents a preliminary miRNA microarray analysis conducted on pooled ascending aorta RNAs obtained from non familial non syndromic TAA patients (five males and five females) compared to matched control pools. Ninety-nine differentially expressed miRNAs with >1.5-fold-up- or down-regulation in TAAs compared to controls were identified, 16.0% of which were similarly regulated in the two sexes.

Genes putatively targeted by differentially expressed miRNAs belonged preferentially to focal adhesion and adherens junction pathways. The results indicate an altered regulation of miRNA-mediated gene expression in the cellular interactions of aneurysmal aortic wall.

Keywords: Adherens junction, Focal adhesion, MicroRNAs (miRNAs), Thoracic aortic aneurysms (TAAs)

[1] Department of Life and Reproduction Sciences, University of Verona, Verona, Italy
[2] Department of Surgery, University of Verona, Verona, Italy

INTRODUCTION

Thoracic aortic aneurysms (TAAs) are characterized by a pathological enlargement of the aorta caused by a maladaptive remodeling of the vessel in response to stress and physiological stimuli. The physiological remodeling process within the aortic vascular wall operates to maintain normal aortic function, whereas pathological remodeling can result in excessive degradation of critical extracellular matrix (ECM) components, leading to the loss of mechanical strength and integrity, aortic dilation, dissection, or rupture.

While more than 20.0% TAAs are inherited as a single gene disorder (*e.g.,* fibrillin-1 gene, FBN1, in Marfan syndrome), the majority of cases are sporadic [1]. The aneurysmal process is now understood to be driven by an unbalanced production of extracellular proteases and inhibitors, but the upstream signalling events are still largely unknown, and especially so for non syndromic events [2]. As recently indicated, an impairment of fine tuning of gene expression in the arterial wall could be related to an altered microRNAs (miRNAs) expression pattern [3].

MicroRNAs are a class of endogenous, small, non coding RNAs regulating the expression of protein-coding genes through pairing with sites in the 3' untranslated region (3'UTR) of their messenger RNA. Due to a perfect or imperfect match, every miRNA may regulate the expression of one

or more genes, and it is likely that at least 30.0% of the genes in a cell may be directly regulated by miRNAs [4]. Recent studies have demonstrated that miRNAs are highly expressed in the vasculature and act as important determinants of disease for the cardiovascular system [5]. In particular, two recent studies in humans have reported altered miRNA expression patterns in TAAs and in thoracic aortic dissection after real time-polymerase chain reaction (ReTi-PCR) or by microarray analysis, respectively [6,7]. Although not yet fully demonstrated, vascular gene expression is thought to be sex related [8-10].

The aim of this study was to compare the miRNA profiles of ascending aortas from TAA patients and controls by microarray analysis. To check for a possible sex effect, different hybridizations were performed for male and female RNA pools.

MATERIALS AND METHODS

Patients and Biological Samples. Samples of aneurysmal ascending aortic wall were obtained from 10 patients (five males and five females) affected by non familial non syndromic TAA, during surgical repair of their ascending aneurysms. The syndromic nature of aortic aneurysm has been systematically excluded by careful evaluation of clinical and family history. Patients with aortic dissections, ruptured aneurysms, Marfan syndrome, or other known connective tissue disorders, were excluded from the study. Autoimmune and/or infectious inflammatory diseases, or chest trauma, were also excluded. Control samples of ascending aorta were obtained from 10 heart transplant recipients without aortic aneurysms (five males and five females). All the individuals had a tricuspid aortic valve. Mean age was 66 ± 9 years. The study conforms with the principles outlined in the Declaration of Helsinki. Aortic samples were promptly dipped into RNAlater solution (Ambion, Austin, TX, USA) in order to preserve their cellular RNA and maintained at room temperature for 2 hours to facilitate liquid permeation. The samples were then stored at –80°C.

Preparation of Microarrays. Total RNA was extracted from the 20 aorta specimens with TRIzol reagent according to the manufacturer's protocol. Total RNA integrity was assessed by an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and RNA integrity numbers (RIN) were sufficient for micro RNA (miRNA) microarray experiments (*i.e.,* RIN >6) [11]. Four samples were then prepared by pooling corresponding RNAs: TAA males, TAA females, control males, and control females.

Sample labeling, PIQOR™ mirExplore microarray hybridization and fluorescence signal detection were performed by Miltenyi Biotec GmbH (MACS Service, Köln, Germany). The TAA and control pools were labeled with Hy5 and Hy3, respectively, and competitively hybridized on the same microarray, separately for the two sexes. Fluorescence signals of the hybridized PIQOR™ Microarrays (Miltenyi Biotec GmbH) were detected using a laser scanner from Agilent (Agilent Technologies).

Image and Data Analysis. Mean signal and mean local background intensities were obtained for each spot of the microarray images using the ImaGene® software (Biodiscovery, Hawthorne, CA, USA). Low-quality spots were flagged and excluded from data analysis. Unflagged spots were analyzed with the PIQOR™ Analyzer software (Miltenyi Biotec GmbH) that allows automated data processing of the raw data text files derived from the ImaGene software. This includes background subtraction to obtain the net signal intensity, data normalization, and calculation of the Hy5/Hy3 ratios. As an additional quality filtering step, only spots/genes that had a signal higher than the 50.0% percentile of the background signal intensities were taken into account for the calculation of the Hy5/Hy3 ratio.

Up- and Down-Regulated MicroRNA and Gene Pathways. Normalized mean Hy5/Hy3 ratios were determined for four replicas per gene. There was a specific detection even for closely related miRNA family members. MicroRNAs that were >1.5-fold-up or down-regulated represented putative candidate miRNAs.

To identify molecular pathways potentially altered by the expression of single or multiple miRNAs, Diana mir- Path Software (Athens, Greece) was used [12]. This web-based application performs an enrichment analysis of multiple miRNA target genes comparing each set of miRNA targets to all known KEGG (Kyoto Encyclopedia of Genes and Genomes, Kyoto, Japan) pathways [13].

RESULTS

The oligonucleotide probes on the microarrays complemented 728 mature miRNAs. Differentially expressed miRNAs ranged from –2.24-fold down regulation to +7.20-fold up regulation, and are listed in Table 1. Out of 728 miRNAs detected, 87 resulted differentially expressed in males (11 down- and 76 up-regulated in TAA) and 28 in females (eight down and 20 up-regulated in TAA), compared to their controls. Five and 11 of these miRNAs were down and up regulated in both male and female

Table 1. Ninety-nine differently expressed microRNAs in ascending aortic aneurysm sex-specific RNA pools. Results are indicated as ratios of TAA pool expression (males and females) to their corresponding control pool expression (males or females); blank boxes represent no differential expression. Up-regulated microRNAs have ratios of >1.5, while down-regulated microRNAs have ratios of <0.66 (if the relative ratio is below 1.0, divide by your ratio to get a fold decrease: 1.0/0.66 = –1.5).

microRNA	Males	Females	microRNA	Males	Females	microRNA	Males	Females
MIR-940	0.48	0.45	**MIR-370**	1.64		**MIR-193A-3P**		2.25
MIR-923	0.56		**HIV1-MIR-H1**	1.64		**MIR-768-5P**	2.27	
MIR-663		0.45	**MIR-383**	1.64		**MIR-376C**	2.29	
MIR-133A	0.56	0.48	**MIR-487B**	1.67	1.70	**MIR-373**	2.30	
MIR-133B	0.56	0.50	**MIR-150**	1.67		**MIR-765**	2.32	
MIR-125A-3P	0.57	0.47	**MIR-340-5P**	1.67		**MIR-493***	2.34	
MIR-486-5P	0.57	0.57	**MIR-183***	1.68		**MIR-422A**	2.35	1.56
MIR-145	0.61		**MIR-324-3P**	1.69		**MIR-23B***	2.37	
MIR-200B*	0.62		**MIR-199A-3P**	1.70		**EBV-MIR-BART13**	2.51	
MIR-193A-5P	0.63		**MIR-29B**	1.74	1.51	**MIR-17***	2.64	
MIR-22	0.63		**MIR-222**		1.75	**MIR-212**		2.46
MIR-665		0.63	**MIR-487A**	1.76		**MIR-373***	2.68	
MIR-638		0.63	**MIR-451**	1.76		**MIR-520E**	3.76	
MIR-193B	0.65		**MIR-210**		1.77	**MIR-15B**	2.79	
MIR-146A	1.51		**MIR-345-5P**	1.79		**MIR-25**	2.80	1.66
MIR-150*	1.52		**MIR-886-5P**	1.81		**MIR-26B**	2.84	
SV40-MIR-S1-5P	1.52		**MIR-874**	1.82		**MIR-187***	2.95	
MIR-23B	1.52		**MIR-30.E**	1.82		**MIR-378**	2.96	
MIR-425	1.54		**MIR-130B**		1.88	**MIR-138-2***	2.98	
MIR-520A-3P	1.54		**MIR-421-3P**	1.95		**MIR-203**	3.00	
MIR-24-2*		1.55	**MIR-520A**	1.97		**MIR-302C**	3.07	
MIR-376A	1.56		**MIR-214***	1.97		**MIR-15A**	3.09	1.85
MIR-557	1.56		**MIR-188-5P**	1.97		**MIR-26A-2***	3.21	
MIR-195	1.56		**MIR-188-3P**		1.97	**MIR-138-1***	3.49	1.82
MIR-628-3P		1.56	**MIR-16-2***	2.00		**MIR-223**	3.53	
MIR-101	1.59		**MIR-16**	2.02		**MIR-148A**	3.55	
MIR-675	1.59		**MIR-654-3P**	2.02		**MIR-221***	3.76	
MIR-491-3P		1.58	**MIR-16-1***	2.04		**MIR-146B-5P**	4.00	2.75
MIR-532-5P	1.60		**MIR-374B**	2.04		**MIR-21***	4.21	2.46
MIR-518F*	1.60		**MIR-34B-3P**	2.13		**MIR-142-5P**	4.42	2.70
MIR-140-5P	1.60		**MIR-744**	2.15		**MIR-21**	4.69	2.71
MIR-513A-5P	1.62		**MIR-185**	2.17		**MIR-128**	4.80	2.13
MIR-216A	1.62		**HIV1-MIR-N367**	2.23		**MIR-126-3P**	7.20	

TAAs, respectively, for a total of 16 sex-insensitive differentially expressed miRNAs. No miRNA was regulated in the opposite direction in one sex compared to the other.

Table 2 shows the most interesting coordinated gene expression pathways resulting from Diana mirPath software analysis. A multiple testing correction was performed according to Bonferroni's method [14], as indicated in Table 2. Focal adhesion was the most gene enriched pathway. It was putatively targeted by the 11 sex-insensitive up-regulated miRNAs (adjusted *p* value 5.43E-09). Adherens junction-related genes were also preferentially indicated, putatively targeted by the five sex-insensitive down-regulated miRNAs (adjusted *p* value 7.12E-06) and apparently also by the up-regulated miRNAs (adjusted *p* value 2.56E-04). Another significant pathway was the regulation of actin cytoskeleton (adjusted *p* value 3.12E-04) that resulted from up-regulated miRNAs analysis.

DISCUSSION

In this study, RNA pools obtained from ascending thoracic aortic wall fragments of patients affected by TAA were competitively hybridized with control pools on microarrays spotted with oligonucleotides putatively recognizing 728 miRNAs. Ninety-nine miRNAs were differentially expressed. The miRNA sequences listed were matched to gene sequences and linked to annotated pathways of gene expression. The most enriched pathways, which included a high number of putative target genes of differentially expressed miRNAs, were focal adhesion and adherens junction.

The most interesting result concerned the 11 up-regulated miRNAs, both in male and female TAAs, which putatively matched 61 genes related to the adhesion processes. This indicates a significant modulation, *i.e.,* repression, of the focal adhesion pathway. The adherens junction pathway, on the basis of the putative targeting by miRNAs, resulted as repressed and stimulated in 26 and 15 genes, respectively. These opposite signals could be a manifestation of the cellular cohesion impairment and of the attempt to reconstitute the integrity of the aortic wall during the development of TAA.

No sex differences were observed in the present study, as miRNAs differentially expressed in males versus females targeted genes belonging to the same general pathways. Some of the differentially expressed miRNAs identified in this study are in agreement with the literature, as reported below.

Down-Regulated microRNAs. Phenotypic abnormalities of vascular smooth muscle cells (VSMCs) and cardiomyocytes have been observed in MIR-133 knockout mice [15]. MIR-145 is decreased in aortas from patients with an aneurysm and was suggested as a potential biomarker for vascular diseases [16].

Up-Regulated microRNAs. MIR-126, the most up-regulated miRNA in this study, has been implicated in the maintenance of vascular integrity [17] and in vascular cell adhesion molecule expression [18]. MIR-29B-mediated down-regulation of

Table 2. Enrichment analysis results of differentially expressed microRNAs (both in male and female TAAs), obtained by Diana mirPath software.

Common MicroRNAs	KEGG Pathways[a]	Number of Genes[b]	*p* Value[c]	Adjusted *p* Value[d]
Down-regulated (5)	adherens junction	15	5.20E-08	7.12E-06
	transforming growth factor β signaling	13	4.35E-04	n.s.
	focal adhesion	21	1.00E-03	n.s.
Up-regulated (11)	focal adhesion	61	3.31E-11	5.43E-09
	adherens junction	26	1.56E-06	2.56E-04
	regulation of actin cytoskeleton	55	1.90E-06	3.12E-04

[a] Pathways with higher enrichment in genes putatively targeted by the miRNAs indicated.
[b] Number of genes belonging to the pathway and putatively targeted by all the miRNAs indicated.
[c] *p* Values calculated by input dataset enrichment analysis performed by a Pearson's chi-squared test as indicated by Diana mirPath algorithm [12].
[d] *p* Values Bonferroni-corrected for multiple testing.

ECM proteins predisposes the aorta to the formation of aneurysms [19]. MIR-21 has been investigated extensively in various tissues and it has been found to promote vascular smooth muscle cell (VSMC) proliferation [16]. Knockdown of MIR-221 and MIR-222 by antisense oligonucleotide miRNA depletion has been foun to reduce VSMCs proliferation in response to vascular injury, and both miRNAs are strongly elevated *in vivo* in VSMCs following vessel injury [16]. MIR-146, MIR-24 and MIR-26 have been implicated in VSMCs proliferation and contraction [16]. MIR-15A, MIR-16, MIR-16-1, MIR-16-2, and MIR-195 belong to a family of miRNAs (the miR-15 family) consistently found to be up-regulated in cardiovascular diseases [5].

In conclusion, and in concordance with other investigators [20], our study indicates that the weakness of the ascending aortic tissue in TAA is linked with a perturbation of cell adhesion and cell interaction gene expression pathways. MicroRNAs probably act as regulators, possibly driving, or at least influencing, the development of the disease.

REFERENCES

1. Keramati AR, Sadeghpour A, Farahani MM, Chandok G, Mani A. The non-syndromic familial thoracic aortic aneurysms and dissections maps to 15q21 locus. BMC Med Genet. 2010;11(11):143-148.
2. Barbour JR, Spinale FG, Ikonomidis JS. Proteinase systems and thoracic aortic aneurysm progression. J Surg Res. 2007; 139(2): 292-307.
3. Qin S, Zhang C. MicroRNAs in vascular disease. (Review) J Cardiovasc Pharmacol. 2011; 57(1): 8-12.
4. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005; 120(1): 15-20.
5. Small EM, Frost RJ, Olson EN. MicroRNAs add a new dimension to cardiovascular disease. (Review) Circulation. 2010; 121(8): 1022-1032.
6. Jones JA, Stroud RE, O'Quinn EC, *et al.* Selective microRNA suppression in human thoracic aneurysms: relationship of miR-29a to aortic size and proteolytic induction. Circ Cardiovasc Genet. 2011; 4(6): 605-613.
7. Liao M, Zou S, Weng J, *et al.* A microRNA profile comparison between thoracic aortic dissection and normal thoracic aorta indicates the potential role of microRNAs in contributing to thoracic aortic dissection pathogenesis. J Vasc Surg. 2011; 53(5): 1341-1349.
8. Sieveking DP, Lim P, Chow RWY, *et al.* A sex-specific role for androgens in angiogenesis. J. Exp Med. 2010; 207(2): 345-352.
9. Arain FA, Kuniyoshi FH, Abdalrhim AD, Miller VM. Sex/gender medicine – the biological basis for personalized care in cardiovascular medicine. Circ J. 2009; 73(10): 1774-1782.
10. Orshal JM, Khalil RA. Gender, sex hormones, and vascular tone. (Review) Am J Physiol Regul Integr Comp Physiol. 2004; 286(2): R233-R249.
11. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. (Review) Mol Aspects Med. 2006; 27(2-3): 126-139.
12. Papadopoulos GL, Alexiou P, Maragkakis M, Reczko M, Hatzigeorgiou AG. DIANA-mirPath: Integrating human and mouse microRNAs in pathways. Bioinformatics. 2009; 25(15): 1991-1993.
13. Kanehisa M, Araki M, Goto S, *et al.* KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008; 36(Database issue): D480-D484.
14. Abdi H. Bonferroni and Šidák corrections for multiple comparisons. In: Salkind NJ, Ed. Encyclopedia of Measurement and Statistics. Thousand Oaks: Sage, 2007.
15. Liu N, Bezprozvannaya S, Shelton JM, *et al.* Mice lacking microRNA 133a develop dynamin 2-dependent centronuclear myopathy. J Clin Invest. 2011; 121(8): 3258-3268.
16. Davis-Dusenbery BN, Wu C, Hata A. Micromanaging vascular smooth muscle cell differentiation and phenotypic modulation. (Review) Arterioscler Thromb Vasc Biol. 2011; 31(11): 2370-2377.
17. Wang S, Aurora AB, Johnson BA, *et al.* The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. Dev Cell. 2008; 15(2): 261-271.
18. Harris TA, Yamakuchi M, Ferlito M, Mendell JT, Lowenstein CJ. MicroRNA-126 regulates endothelial expression of vascular cell adhesion molecule 1. Proc Natl Acad Sci USA. 2008; 105(5): 1516-1521.
19. Boon RA, Seeger T, Heydt S, *et al.* MicroRNA-29 in aortic dilation: implications for aneurysm formation. Circ Res. 2011; 109(10): 1115-1119.
20. Kim JH, Na CY, Choi SY, *et al.* Integration of gene-expression profiles and pathway analysis in ascending thoracic aortic aneurysms. Ann Vasc Surg. 2010; 24(4): 538-549.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

GENETICS OF NON SYNDROMIC HEARING LOSS IN THE REPUBLIC OF MACEDONIA

Sukarova Stefanovska E[1], Cakar M[2], Filipce I[2], Plaseska Karanfilska D[1]

***Corresponding Author:** Emilija Sukarova Stefanovska, Ph.D., Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov", Macedonian Academy of Sciences and Arts, Av. Krste Misirkov 2, POB 428, 1000 Skopje, Republic of Macedonia; Tel.: +389(0)2-3235-410; Fax: +389-(0)2-3115-434; E-mail: emilija@manu.edu.mk

ABSTRACT

Hearing impairment is the most common sensory deficit in humans affecting 1 in 1000 newborns. When present in an infant, deafness may have dramatic effects on language acquisition, seriously compromising the quality of their life. Deafness is influenced by both genetic and environmental factors, with inherited causes as the most prominent etiological factor in deafness in developed countries. The genetic basis of hearing loss is complex with numerous loci and genes underlying hereditary sensoryneural non syndromic hearing loss (NSHL) in humans. Despite the wide functional heterogeneity of the genes, mutations in the *GJB2* gene are found to be the most common cause of sporadic and recessive NSHL in many populations worldwide. Molecular characterization of deafness in the Republic of Macedonia was performed in 130 NSHL profoundly deaf children from different ethnic origins. Molecular studies included direct sequencing of the *GJB2* gene and specific polymerase chain reaction (PCR) analyses for the del(*GJB6*-D13S1830) mutation. Five common mitochondrial DNA (mtDNA) mutations [A1555G, 961delT+ C(n), T1095C, C1494T and A827G] were also analyzed using the SNaPShot method. In preliminary studies, *GJB2* gene mutations were found in 36.4% of analyzed patients, with predominance of 35delG in Macedonian and Albanian patients and W24X in Gypsy patients, respectively. No del(*GJB6*-D13S1830) mutation was found. None of the analyzed deafness-associated mutations in mtDNA were identified in the studied patients.

Keywords: DFNB1 locus, *GJB2* gene mutations, Non syndromic hearing loss (NSHL).

INTRODUCTION

Hearing impairment is the most common sensory deficit in humans. Approximately 1 in 1000 newborns is affected by severe to profound deafness at birth, and additionally 1 in 1000 children progressively develops hearing loss during childhood. With the age, hearing loss progressively increases, so that over 50.0% of individuals above 80 years of age are affected [1]. The World Health Organization (WHO) has estimated that in 2005, 278 million people worldwide have moderate to profound hearing loss (WHO; http://www.who.int/mediacentre/factsheets/fs300/ en/index.html). In the Republic of Macedonia, there are 9000 deaf individuals. When present in an infant, deafness may have dramatic effect on language acquisition, seriously compromising the quality of their life.

[1] Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Skopje, Republic of Macedonia

[2] Othorinolaryngology Clininc, Clinical Center University Hospital, Skopje, Republic of Macedonia

Deafness is influenced by both genetic and environmental factors, with inherited causes defined as the most prominent etiological factor in deafness in developed countries. When associated with other symptoms it is referred to as syndromyc hearing loss (SHL), while when it occurs as an isolated symptom, it is referred to as non syndromic hearing loss (NSHL). Non syndromic hearing loss accounts for about 70.0% of genetic deafness and is almost exclusively monogenic and highly heterogeneous [2,3]. Non syndromic hearing loss can be transmitted as autosomal recessive (80.0%), autosomal dominant (20.0%), X-linked deafness and in mitochondrial mode (2.0%). The autosomal recessive forms of deafness are generally the most severe and are almost exclusively caused by cochlear defects of hair cells (sensoryneural deafness). Autosomal dominant forms of deafness are usually post-lingual and progressive.

The genetic basis of hearing loss is complex. At least 1.0% of human protein-coding genes are involved in sound perception. There are approximately 113 mapped loci and 51 different genes that have been shown to underlie hereditary sensoryneural NSHL in humans (http:webhost.ua.ac. be/hhh/). They encode a wide variety of protein classes such as transcription factors, cytoskeletal and extracellular matrix components, and ion channels. Mutations in different genes can lead to the same clinical manifestation of deafness. Alternatively, different mutations in the same gene may result in syndromic, non syndromic, recessive or dominant deafness; finally, the same mutation can be associated with quite variable phenotypes [4].

Despite the wide functional heterogeneity of the genes involved in the perception of sound, mutations in the DFNB1 locus on chromosome 13 (13q12) are responsible for more than half of all cases of NSHL. The locus contains three genes, *GJB2, GJB6* and *GJA1*, encoding for the transmembrane gap junction proteins connexin 26, connexin 30 and connexin 31, respectively, involved in ion transfer and homeostasis in the inner ear [5,6].

Mutations in the *GJB2* gene are the most common cause of sporadic and recessive NSHL, in many populations worldwide [5]. More than 100 different mutations in this gene have been described with specific prevalence in different ethnic groups and geographic regions. In Caucasians, the most common mutation is 35delG [7], 167delT in Ashkenazi Jews [8], 235delC in Japanese [9], while W24X is frequent in India and in Gypsies. Due to the high incidence of *GJB2* gene mutations, molecular testing for *GJB2* gene mutations has rapidly become the standard of care for the diagnosis and counseling of patients with non syndromic hearing impairment of unknown cause.

Molecular Characterization of Deafness in the Republic of Macedonia. Molecular characterization of deafness in the Republic of Macedonia started in 2006, as a project of the Macedonian Academy of Sciences and Arts (MASA) in collaboration with the Audiology Centre at the Clinic of Othorinolaryngology, Medical Faculty, Skopje. Until now, 130 unrelated cases of different ethnic origins [Macedonians (78), Albanians (19), Gypsies (30) and Turks (3)], all with profound NSHL, were ascertained through the Audiology Centre, where audiological examinations and detailed family history analyses were performed.

***GJB2* Gene Mutation Analysis.** All patients and members of their families were screened for the presence of the *GJB2* gene mutations using single strand conformation polymorphism analysis, restriction enzyme digestion or direct sequencing of the coding exon 2. Multiplex ligation-dependent probe amplification (MLPA) analysis using SALSA MLPA kit P163-C1 Hearing loss probemix (MRC Holland, Amsterdam, The Netherlands) was performed in order to determine the presence of deletions/duplications in the 13q region, as well as the specific point mutations in the *GJB2* gene (splice site mutation IVS1+1G>A, 35delG, 101T>C, 167delT, 235delC and 313del14).

Preliminary studies on mutations in the *GJB2* gene in 33 persons with pre lingual non syndromic deafness in the Republic of Macedonia determined a prevalence of 36.4%, with 35delG as the most frequent variation found in 68.2% of mutated chromosomes, followed by W24X (18.2%), predominantly in Gypsy families, V37I (9.1%) and R127H (4.5%) [10]. Additionally two other mutations were found in the analyzed patients, Cd120delGAG and T175I, each with a frequency of 0.8%. None of the analyzed patients carry the IVS-I-1 (G>A) mutation. These findings indicate that as in other Caucasian populations, the 35delG mutation is the most frequent cause of deafness in the Republic of Macedonia and

should be tested in each routine diagnostic approach in the Macedonian population, while the W24X mutation is the first mutation that should be tested in the deaf Macedonian Gypsy population.

Another member of the connexin family, connexin 30, is encoded by the *GJB6* gene located 350 kb upstream of *GJB2* within the same DFNB1 locus. One large deletion, del(*GJB6*-D13S1830), has been reported as a common cause of non syndromic deafness, truncating the *GJB6* gene and also abolishing *GJB2* expression, possibly by deleting a currently unidentified *GJB2* regulatory element [11]. This deletion, either in the homozygous state or in combination with a *GJB2* gene mutation, has been described as the cause of non syndromic deafness at 5.0-9.0% of all DFNB1 alleles in Europe [12]. Screening for the presence of del(*GJB6*-D13S1830) in the *GJB6* gene revealed no mutation in our group of analyzed patients.

Mitochondrial DNA Mutations Analysis. Mutations in mitochondrial DNA (mtDNA) are found to contribute to sensoryneural deafness, including both syndromic and non syndromic forms, with a variable frequency in different populations of deaf persons [13]. Hot-spot regions for deafness mutations are the *MT-RNR1* gene, encoding the 12S ribosomal RNA (rRNA) and the *MT-TS1* gene, encoding the mitochondrial transfer RNA (tRNA) serine 1. One hundred and thirty Macedonian patients with NSHL were screened for the five mitochondrial mutations associated with deafness (A827G, 961delT+Cn, T1095C, C1494T and A1555G) by a single nucleotide primer extension assay utilizing the ABI PRISM™ SNaPshot Multiplex Kit (Life Technologies, Carlsbad, CA, USA) following the manufacturer's instructions [14]. None of the analyzed deafness-associated mutations were identified in the studied patients.

These findings suggest that mtDNA mutations do not represent risk factor for sensoryneural deafness in the Macedonian population.

REFERENCES

1. Cohen MM, Gorlin RJ. Epidemiology, etiology and genetic patterns. In: Gorlin RJ, Toriello HV, Cohen MM, Eds. Hereditary Hearing Loss and Its Syndromes. Oxford: Academic Press. 1995: 9-21.
2. Schrijver I. Hereditary non-syndromic sensorineural hearing loss. Transforming silence to sound. J Molec Diagn. 2004; 6(4): 275-284.
3. Petit C. From deafness genes to hearing mechanisms: harmony and counterpoint. Trends Mol Med. 2006; 12(2): 57-64.
4. Mese G, Londin E, Mui R, Brink PR, White TW. Altered gating properties of functional Cx26 mutants associated with recessive non-syndromic hearing loss. Hum Genet. 2004; 115(3): 191-199.
5. Rabionet R, Zelante L, Lopez-Bigas N, *et al.* Molecular basis of childhood deafness resulting from mutations in the GJB2 (Connexin 26) gene. Hum Genet. 2000; 106(1): 40-44.
6. Kennenson A, Van Naarden Brown K, Boyle C. GJB2 (Connexin 26) variants and nonsyndromic sensorineural hearing loss: A HuGE review. Genet Med. 2002; 4(4): 258-274.
7. Zelante I, Gasparini P, Estivill X, *et al.* Connexin 26 mutations associated with the most common form of non-syndromic neurosensory autosomal recessive deafness (DFNB1) in Mediterraneans. Hum Mol Genet. 1997; 6(9): 1605-1609.
8. Morell RJ, Kim HJ, Hood LJ, *et al.* Mutations in the connexin 26 gene (GJB2) among Ashkenazi Jews with nonsyndromic recessive deafness. N Engl J Med. 1998; 339(21): 1500-1505.
9. Abe S, Kelly PM, Kimberling WJ, Usami SI. Connexin 26 gene (GJB2) mutation modulates the severity of hearing loss associated with the 1555A>G mitochondrial mutation. Am J Med Genet. 2001; 103(4): 334-338.
10. Sukarova Stefanovska E, Momirovska A, Cakar M, Efremov GD. GJB2 mutations in non syndromic hearing loss in the Republic of Macedonia. Balkan J Med Genet. 2009; 12(2): 11-17.
11. Del Castillo I, Villamar M, Moreno-Pelayo MA, *et al.* A deletion involving the connexion 30 gene in nonsyndromic hearing impairment. N Engl J Med. 2002; 346(4): 243-249.
12. Del Castillo I, Moreno-Pelayo MA, Del Castillo FJ, *et al.* Prevalence and evolutionary origins of the del(GJB6-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: a multi-center study. Am J Hum Genet. 2003; 73(6): 1452-1458.
13. Maniglia LP, Bruna Moreira CL, Menezes da Silva MA, Piatto VB, Maniglia JV. Screening of the mitochondrial A1555G mutation in patients with sensorineural hearing loss. Rev Bras Otorinolaryngol. 2008; 74(5): 731-736.
14. Bardien S, Human H, Harris T, *et al.* A rapid method for detection of five known mutations associated with aminoglycoside-induced deafness. BMC Med Genet. 2009; 10: 2.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

MOLECULAR DIAGNOSTICS OF β-THALASSEMIA

Atanasovska B[1], Bozhinovski G[1], Chakalova L[1], Kocheva S[2], Karanfilski O[3], Plaseska-Karanfiska D[1,*]

***Corresponding Author:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov", Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

ABSTRACT

A high-quality hemoglobinopathy diagnosis is based on the results of a number of tests including assays for molecular identification of causative mutations. We describe the current diagnostic strategy for the identification of β-thalassemias and hemoglobin (Hb) variants at the International Reference Laboratory for Haemoglobinopathies, Research Centre for Genetic Engineering and Biotechnology (RCGEB) "Georgi D. Efremov," Skopje, Republic of Macedonia. Our overall approach and most of the methods we use for detection of mutations are designed for the specific target population. We discuss new technical improvements that have allowed us to substantially reduce the average time necessary for reaching a conclusive diagnosis.

Keywords: Hb Lepore; Hemoglobinopathy; Molecular diagnostics; Thalassemia.

[1] Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Skopje, Republic of Macedonia

[2] Pediatric Clinic, Medical Faculty, University "Stt. Cyril and Methodius," Skopje, Republic of Macedonia

[3] Hematology Clinic, Medical Faculty, University "Stt. Cyril and Methodius," Skopje, Republic of Macedonia

INTRODUCTION

Hemoglobinopathies are caused by genetic defects affecting the globin genes encoding for the α and β chains of the hemoglobin (Hb) molecule. In the Mediterranean region in particular, there is a high incidence of mutations disturbing the function of the *HBB* gene [1]. Some of the mutations reduce or eliminate the expression of the *HBB* gene leading to net Hb deficiency and β-thalassemia (β-thal) [2]. Other mutations give rise to abnormal Hb variants such as Hb S [β6(A3)Glu→Val, G*A*G>G*T*G] and Hb Lepore-Boston-Washington (Hb LBW; δ87-βIVS-II-8) [2-4]. The severity of the clinical symptoms depends on the molecular consequences of the genetic abnormality or combination thereof and is modulated by other genetic and environmental factors [2,5,6].

The best practice in hemoglobinopathy diagnostics involves molecular identification of the causative mutations. Molecular diagnostics of β-thal has been a major focal point of the activities at the International Reference Laboratory for Haemoglobinopathies, Research Centre for Genetic Engineering and Biotechnology (RCGEB) "Georgi D. Efremov", Skopje, Republic of Macedonia at Skopje, Republic of Macedonia. Thousands of cases have been tested for the presence of mutations throughout the years. This study has helped determine the frequencies of the most common

mutations in Macedonia and several neighboring countries [7-10]. Until recently, the primary method for the detection of β-thal mutations in our laboratory was allele-specific oligonucleotide hybridization, whereby consecutive rounds of hybridization with several mutation-specific probes were performed. We were looking to reduce the time necessary for reaching a definitive diagnosis by introducing a semi-automated technique allowing simultaneous detection of the most commonly occurring β-thal mutations. Unfortunately, published techniques, such as primer extension and melting curve analysis were suboptimal in terms of precision and multiplexing [11-13]. We therefore set out to develop a new assay for the identification of common Mediterranean mutations that occur at high frequencies in the Republic of Macedonia, namely HBB:c.93-21G>A, HBB:c.92+1G>A, HBB:c.92+6T>C, HBB:c.118C>T, HBB:c.316-106C>G, HBB:c.17_18 del CT, HBB:c.20delA, HBB:c.25_26delAA and HBB:c.20 A>T [14,15]. Our protocol utilizes single-nucleotide primer extension to interrogate the whole panel of mutations in a single, internally controlled reaction. We also designed a complementary duplex polymerase chain reaction (PCR) assay for detection of the Hb LBW deletion NG_000007.3:g.63632_71046 del, the most common Hb variant in our geographic area [9]. Both assays have been thoroughly validated and evaluated as described elsewhere (Atanasovska *et al.,* in preparation). Here we report the overall diagnostic strategy for the identification of β-thal mutations and Hb variants that includes the new assays.

MATERIALS AND METHODS

The human subjects that participated in this study were referred to the RCGEB for laboratory investigations for suspected haemoglobinopathies. Informed consent was obtained in accordance with the Declaration of Helsinki. Peripheral blood samples were obtained by standard venipuncture. Standard blood test results were obtained for each patient. Levels of Hb A, Hb A_2, Hb F and Hb variants were measured by cation exchange high performance liquid chromatography (HPLC) [16]; red cell osmotic fragility was assessed by the osmotic fragility test as described before [17]. Genomic DNA was isolated following a standard phenol extraction/ethanol precipitation protocol. DNA was dissolved in TE buffer, pH 8.0 and stored at 4°C. The multiplex single-nucleotide primer extension assay for detection of Mediterranean β-thal mutations and the duplex PCR assay for detection of the Hb LBW deletion were performed as described elsewhere (Atanasovska *et al.,* in preparation). A subset of samples were analyzed by sequencing of the PCR-amplified *HBB* gene using the ABI PRISM™ Big Dye Terminator v.1.1 Kit (Life Technologies, Carlsbad, CA, USA). The sequences of the primers used for PCR amplification and sequencing are available upon request.

RESULTS AND DISCUSSION

The vast majority of subjects referred to the laboratory were heterozygous carriers of hemoglobinopathy mutations and non carriers referred by their medical specialist or general practitioner to exclude a hemoglobinopathy. Our diagnostic algorithm is presented in Figure 1. When a case is referred to the laboratory for hemoglobinopathy examination, we initially review the patient data including standard blood test results and family history and run a second tier of hematology tests, namely HPLC quantification of normal and abnormal Hb variants and assessment of red cell osmotic fragility. If a Hb variant, *e.g.,* the relatively common Hb LBW, is present in a sample it is identified on the HPLC profile at this stage by virtue of its specific retention time, proportion of total Hb and peak characteristics. Samples identified as Hb LBW heterozygotes or homozygotes by HPLC are tested by the Lepore PCR assay to confirm the presence of the Hb LBW chromosome. This approach warrants unequivocal identification of Hb LBW cases at an early stage of the procedure. These samples are then excluded from further analysis unless a severe clinical picture suggests compound heterozygosity (see below).

Thalassemia is diagnosed based on red blood cell indices combined with the results of the osmotic fragility test and the HPLC analysis. The main diagnostic parameters pointing to probable β-thal trait are: elevated Hb A_2, low total Hb level, low mean corpuscular volume (MCV), low mean corpuscular Hb (MCH), elevated Hb F, decreased osmotic fragility. β-Thalassemia symptoms can vary between carriers, *e.g.,* an individual could feature most if

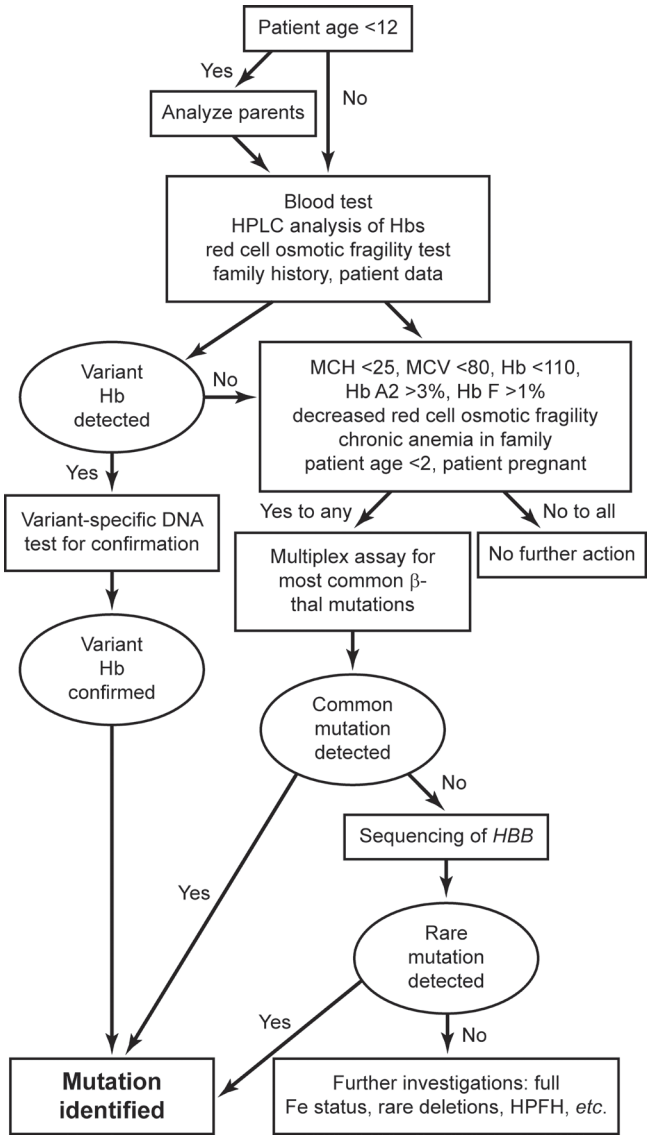


Figure 1. Flow chart of the diagnostic algorithm for identification of patients carrying *HBB* mutations.

not all of these indicators to almost none or borderline values. Using several independent parameters minimizes the risk of missing β-thal carriers in the initial screen. The clinical picture for homozygotes and compound heterozygotes is largely clearer and there is a much lower risk of misdiagnosing these cases.

Based on the results of these analyses. we assign cases for molecular detection of β-thal trait. These samples are first tested for the presence of the eight most common β-thal mutations by the multiplex single-nucleotide primer extension assay. In the past, our protocol for molecular characterization of the *HBB* gene was prohibitively time-consuming, forcing us to apply relatively stringent inclusion criteria. The introduction of the new multiplex assay allows us to test the majority of the subjects referred to the laboratory. In particular, children are tested for the most common mutations by default even if their blood test results are compatible with a normal genotype. Parents are also invited to provide blood samples so that the diagnosis is cross-checked independently. It has to be pointed out that the presence of high levels of Hb F in children younger than 1 year of age can mask the manifestation of β-thal. It is therefore important to apply definitive DNA tests to eliminate false negative results. Collectively, the most common β-thal mutations detected by the multiplex assay account for approximately 90.0% of all hemoglobinopathy cases in the Republic of Macedonia [7,8]. In case the multiplex assay yields a normal genotype while the hematology data points to the presence of a β-thal mutation, the sample is assigned for sequencing of the *HBB* gene in order to reveal genetic variations not tested in the multiplex assay.

Since we incorporated the multiplex assay and the Lepore PCR assay into the routine hemoglobinopathy work-up in our laboratory, we have processed a total of 186 patient samples. For 83 cases, the data from the various hematological tests were concordant and compatible with a normal genotype and further testing was not necessary. Nevertheless, these samples were assayed by the multiplex assay partly to corroborate the absence of common mutations, partly to assess the assay reproducibility. Hemoglobinopathies were thus excluded in these cases. For the remaining 103 cases, there were indications for genetic abnormalities affecting the *HBB* gene. Three of these patients were confirmed to be Hb LBW heterozygotes (Table 1) and were not tested further. We applied the multiplex thalassemia assay to the remaining 100 samples and identified 95 β-thal heterozygotes, each carrying one of the β-thal mutations included in the assay (Table 1). The multiplex assay failed to identify any mutations in five samples. These were subjected to direct sequencing of the *HBB* gene and were found to carry other β-thal mutations (Table 1) in unison with the hematology data. Thus, a conclusive diagnosis was reached for every case. Importantly, the newly developed procedures have significantly reduced the time and cost necessary to complete the analyses.

Table 1. Mutations detected by the multiplex single-nucleotide extension assay, the Lepore polymerase chain reaction assay and direct sequencing of the *HBB* gene.

HGVS Nomenclature[a]	Mutation[b]	Type	Assay	Chromosomes Detected
HBB:c.93-21G>A	IVS-I-110 (G>A)	β^+	Multiplex	37
HBB:c.92+1G>A	IVS-I-1 (G>A)	β^0	Multiplex	11
HBB:c.92+6T>C	IVS-I-6 (T>C)	β^+	Multiplex	15
HBB:c.118C>T	codon 39 (C>T)	β^0	Multiplex	16
HBB:c.316-106C>G	IVS-II-745 (C>G)	β^+	Multiplex	6
HBB:c.17_18delCT	codon 5 (–CT); CCT(Pro)>C– –	β^0	Multiplex	3
HBB:c.20delA	codon 6 (–A); GAG(Lys)>G–G	β^0	Multiplex	5
HBB:c.25-26delAA	codon 8 (–AA); AAG(Lys)>– –A	β^0	Multiplex	2
NG_000007.3:g.63632_71046del	Hb LBW	Hb+β-thal	Lepore PCR	3
HBB:c.–80T>A	–30 (T>A)	β^+	Sequencing	2
HBB:c.316-3C>A	IVS-II-848 (C>A)	β^+	Sequencing	2
HBB:c.250delG	codons 82/83 (–G); AAG GGC(Lys Gly)> AAG –GC	β^0	Sequencing	1

[a] Patrinos *et al.* [15].
[b] Huisman *et al.* [14].

ACKNOWLEDGMENTS

We are indebted to the late Professor Dr. Georgi D. Efremov for his long-standing support and dedication to hemoglobinopathy research. We are grateful to Dr. Katarina Davalieva, Ms. Ivana Maleva and Dr. Svetlana Madjunkova (RCGEB "Georgi D. Efremov," MASA, Skopje, Republic of Macedonia) for critical reading of the manuscript. We also thank-Ms. Stana Janeva (RCGEB "Georgi D. Efremov," MASA, Skopje, Republic of Macedonia) for technical assistance.

REFERENCES

1. Henderson S, Timbs A, McCarthy J, *et al.* Incidence of haemoglobinopathies in various populations – the impact of immigration. Clin Biochem. 2009; 42(18): 1745-1756.
2. Weatherall DJ. Disorders of the synthesis or function of haemoglobin. In: Warrell DA, Cox TM, Firth JD, Benz EJ Jr, Eds. Oxford Textbook of Medicine, 4th ed. Oxford: Oxford University Press. 2003; 676-677.
3. Baglioni C. The fusion of two peptide chains in Hemoglobin Lepore and its interpretation as a genetic deletion. Proc Natl Acad Sci USA. 1962; 48(11): 1880-1886.
4. Ramirez F, Mears JG, Nudel U, *et al.* Defects in DNA and globin messenger RNA in homozygotes for Hemoglobin Lepore. J Clin Invest. 1979; 63(4): 736-742.
5. Galanello R, Origa R. β-Thalassemia. Orphanet J Rare Dis. 2010;5:11.
6. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. Hum Mol Genet. 2009; 18(R2):R216-R223.
7. Efremov DG. Thalassemias and other hemoglobinopathies in the Republic of Macedonia. Hemoglobin. 2007,31(1):1-15.
8. Dimovski A, Efremov DG, Jankovic L, *et al.* β-Thalassemia in Yugoslavia. Hemoglobin. 1990; 14(1): 15-24.
9. Efremov GD. β-, δβ-Thalassemia and Hb Lepore among Yugoslav, Bulgarian, Turkish and Albanian. Haematologica. 1990; 75(Suppl 5): 31-41.
10. Petkov GH, Efremov GD, Efremov DG, *et al.* β-Thalassemia in Bulgaria. Hemoglobin. 1990; 14(1): 25-33.
11. Kobayashi M, Rappaport E, Blasband A, *et al.* Fluorescence-based DNA minisequence analysis for detection of known single-base changes in genomic DNA. Mol Cell Probes. 1995; 9(3): 175-182.
12. Galbiati S, Chiari M, Macellari M, Ferrari M, Cremonesi L, Cretich M. High-throughput mutational screening for β-thalassemia by singlenucleotide extension. Electrophoresis. 2007; 28(23): 4289-4294.

13. Vrettou C, Traeger-Synodinos J, Tzetis M, Malamis G, Kanavakis E. Rapid screening of multiple β-globin gene mutations by real-time PCR on the LightCycler: application to carrier screening and prenatal diagnosis of thalassemia syndromes. Clin Chem. 2003; 49(5): 769-776.
14. Huisman THJ, Carver MFH, Baysal E. A Syllabus of Thalassemia Mutations (1997). Augusta: The Sickle Cell Anemia Foundation, 1997.
15. Patrinos GP, Giardine B, Riemer C, *et al.* Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. Nucleic Acids Res. 2004; 32(Suppl 1): D537-D541.
16. Samperi P, Mancuso GR, Dibenedetto SP, Di Cataldo A, Ragusa R, Schilirò G. High performance liquid chromatography (HPLC): a simple method to quantify Hb C, O-Arab, Agenogi and F. Clin Lab Haematol. 1991; 13(2): 169-75.
17. Malamos B, Fessas P, Stamatoyannopoulos G. Types of thalassaemia trait carriers as revealed by a study of their incidence in Greece. Br J Haematol. 1964; 8(1): 5-14.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

STUDY OF THE HEPATITIS C VIRUS IN THE REPUBLIC OF MACEDONIA

Kiprijanovska S[1], Sukarova-Stefanovska E[1], Noveski P[1],
Chalovska V[2], Polenakovic M[1], Plaseska-Karanfilska D[1,*]

***Corresponding Author:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov", Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

ABSTRACT

Hepatitis C virus (HCV) is a major public health problem. It is a leading cause of chronic liver disease and the most common indication for liver transplantation. The therapy for eradication of HCV infection is successful in only 50.0-80.0% of patients and is highly dependent on the HCV genotype.

Molecular detection and characterization of HCV in the Republic of Macedonia started in 1990. Since then, more than 4000 samples have been analyzed at the Research Centre for Genetic Engineering and Biotechnology (RCGEB) "Georgi D. Efremov," Skopje, Republic of Macedonia. The prevalence of HCV infections in the healthy population of the Republic of Macedonia was found to be 0.4%, while it varies between 23.0 and 43.0% in different at-risk groups of patients.

The prevalence of HCV genotypes, according to associated risk factors in HCV infected patients from the Republic of Macedonia, was analyzed. We found genotype 1 to be predominant in a group of hemodialysis patients, while genotype 3 was predominant in intravenous (IV) drug users.

Association of six polymorphisms in the Oligoadenylate synthetase (*OASL*)-like interferon-stimulated gene with a sustained virological response was also analyzed. Our preliminary results suggest that non ancestral alleles in four of the six studies polymorphisms in *OASL* gene are associated with sustained virological response among HCV infected patients in R. Macedonia.

Keywords: Hepatitis C virus (HCV) Genotyping, Intravenous (IV) drug users, HCV therapy, Oligoadenylate synthetase (*OASL*) like interferon stimulated gene

INTRODUCTION

Hepatitis C virus (HCV) is a major health problem affecting 170 million people worldwide. The prevalence rate is about 1.0% in western countries and North America, 3.0-4.0% in some Mediterranean and Asian countries and up to 10.0-20.0% in parts of central Africa [1]. The HCV infection is the leading cause of chronic hepatitis worldwide, progressing to liver cirrhosis, and hepatocellular carcinoma in approximately 20.0% of patients [2-5].

Hepatitis C virus was the first virus discovered by molecular cloning [6]. Hepatitis C virus is an enveloped, positive-sense RNA virus, belonging to the Hepacivirus genus of the Flaviviridae family [6]. The genome is approximately 9.6 kb in size

[1] Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Skopje, Republic of Macedonia

[2] Clinic of Gastroenterohepatology, Medical Faculty, Skopje, Republic of Macedonia

and contains an open reading frame (ORF) encoding a large protein precursor [7]. This precursor is cleaved by host and viral proteases into various structural and non structural proteins. The structural proteins are at the 5' end and include the capsid or core protein (c), two envelope proteins (E1 and E2) and a small protein of unknown function (P7). The structural proteins are followed by at least six non structural (NS) proteins denoted as NS2, NS3, NS4A, NS4B, NS5A and NS5B [8].

On the basis of phylogenetic analysis of nucleotide sequences, HCV can be divided into six genotypes and several subtypes. The six genotypes differ between each other in 30.0-35.0% of sequence, while the subtypes differ in 20.0-25.0% over the complete genome. Different HCV genotypes exhibit different epidemiological and clinical implications. The HCV genotypes 1, 2, and 3 appear to have a worldwide distribution. The HCV genotypes 1 and 3 are the most common genotypes in the United States [9] and in Europe [10,11]. Genotype 3, which is endemic in Southeast Asia, has been encountered in Europe and the USA, with relatively high frequency in intravenous (IV) drug users [12]. The HCV genotype 4 appears to be prevalent in North Africa and the Middle East [13,14], and genotype 5 predominates in South Africa. Genotype 6, according to different authors, is divided into additional genotypes 7-11 and is mainly found in the HCV populations of Vietnam and Hong Kong [15-19]. Hepatitis C virus genotypes have proved to be important epidemiologic marker that can be used to predict success of therapy. The HCV genotype 1 is a more aggressive strain and one that is less likely to respond to interferon treatment than HCV genotypes 2 or 3. All this information has great significance when planning future strategies for eradication and therapeutic management of HCV.

Molecular detection and characterization of the HCV infections in the Republic of Macedonia started in 1990. Since then, more than 4000 samples have been analyzed at the Research Centre for Genetic Engineering and Biotechnology (RCGEB) "Georgi D. Efremov," Skopje, Republic of Macedonia. Blood samples were collected by the Clinic of Infection Diseases, Clinic of Gastroenterology and several dialysis centers from the Republic of Macedonia. Amplicor Specimen Preparation and Amplification kit (Roche Diagnostics, Indianapolis, IN, USA) were used for RNA isolation and amplification according to the manufacturer's recommendations. The prevalence of HCV infections in the healthy population of the Republic of Macedonia was found to be 0.4%, while in different at-risk groups of patients such as IV drug users, hemodialysis patients, patients under a blood transfusion regimen and those with unknown factors, the prevalence of HCV varies between 23.0 and 43.0% [20].

Genotyping analyses were performed on 1346 patients with a positive HCV-RNA analysis result. The HCV/RNA genotyping was performed with an in-house allele-specific oligonucleotide (ASO) hybridization method using specific oligonucleotide probes for different HCV genotypes [genotype 1 (5'-CGC TCA ATG CCT GGA GAT-3'); HCV2a (5'-CAC TCT ATG CCC GGC CAT-3'), HCV2b (5'-CAC TCT ATA CCC GGC CAT-3'), HCV3 (5'-CGC TCA ATA CCC AGA AAT-3') and HCV4a (5'-CAC TCT ATG CCC GGC C-3)].

Genotypes 1 and 3 are predominant in patients from the Republic of Macedonia. Genotype 1 was found in 55.4% of genotyped patients (n = 714), while genotype 3 was found in 44.6% of genotyped patients (n = 632) [20]. There was statistically significant difference between the risk factors analyzed and the acquisition of HCV infection.

Hepatitis C virus is a highly prevalent infection in chronic dialysis patients (37.7%) and represents one of the major problems of hemodialysis units in our country [21]. Genotype 1 is predominant in this group of patients (90.7%). A combined infection (4.6%) of genotypes 2 and 3 was found in two patients (4.6%), and a combined infection of genotypes 1 and 3 was also found in another two patients (4.6%) [22].

We found predominance of HCV genotype 3 in HCV-positive IV drug users (93.35%). High prevalence of genotype 3 in an analyzed group of IV drug users is similar to the pattern of genotypes in IV drug users in both Europe and the USA [12]. The HCV genotype distribution in our patients over the years, shows a shift between the prevalence of genotypes 1 and 3. The increase of patients with genotype 3 is due to the increasing number of IV drug users analyzed at our center.

Evidence from several studies indicate that interferon signaling pathway genes *(IPGs*) and interferon stimulated genes *(ISGs*) are associated with

the host response to HCV infection. In order to investigate the possible association of six polymorphisms in the *OASL*-like interferon stimulated gene with a sustained virological response, we studied six single nucleotide polymorphisms in the *OASL* gene in two groups of patients with an HCV infection: patients who were non responders to the therapy, and those who had sustained a virological response to the therapy. Our preliminary results suggest that non ancestral alleles in four of the six studied polymorphisms in the *OASL* gene are associated with sustained virological response to pegylated interferon α-2a (Peg-IFN-α) plus ribavirin therapy [23].

REFERENCES

1. World Health Organization (WHO). Global surveillance and control of hepatitis C. Report of a WHO Consultation organized in collaboration with the Viral Hepatitis Prevention Board, Antwerp, Belgium. J Viral Hepat. 2001; 6:35-47.
2. Seeff LB. Natural history of viral hepatitis, type C. Semin Gastrointest Dis. 1995; 6(1): 20-27.
3. Di Bisceglie AM. Hepatitis C and hepatocellular carcinoma Semin Liver Dis. 1995; 15(1): 64-69.
4. Kuo G, Choo QL, Alter HJ, *et al.* An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. Science. 1989; 244(4902): 362-364.
5. Lauer GM, Walker BD. Hepatitis C virus infection. N Engl J Med. 2001; 345(1): 41-52.
6. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science. 1989; 244(4902): 359-362.
7. Houghton M. Hepatitis C virus, In: Fields BN, Knipe DM, Howley PM, Chanock RH, Eds. Fields Virology, 3rd ed. Philadelphia: Lippincott-Raven. 1996: 1035-1058.
8. Al-Faleh F, Ramia S. Hepatitis C virus infection in Saudi Arabia: a review. Ann Saudi Med. 1997; 17(1): 77-82.
9. Bhattacherjee V, Prescott LE, Pike I, *et al.* Use of NS-4 peptides to identify type-specific antibody to genotypes 1, 2, 3, 4, 5 and 6. J Gen Virol. 1995; 76(Pt 7): 1737-1748.
10. McOmish F, Yap PL, Dow BC, *et al.* Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. J Clin Microbiol. 1994; 32(4): 884-892.0
11. Mondelli MU, Silini E. Significance of hepatitis C virus genotypes. Viral Hepat Rev. 1995; 1(Suppl 1): 111-120.
12. Silini E, Bono F, Cividini A, *et al.* Molecular epidemiology of hepatitis C virus infection among intravenous drug users. J Hepatol. 1995; 22(6): 691-695.
13. Abdulkarim AS, Zein NN, Germer JJ, *et al.* Hepatitis C virus genotypes and hepatitis G virus in hemodialysis patients from Syria: identification of two novel hepatitis C virus subtypes. Am J Trop Med Hyg. 1998; 59(4): 571-576.
14. Chamberlain RW, Adams N, Saeed AA, Simmonds P, Elliot RM. Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the Middle East. J Gen Virol. 1994; 78(Pt 6):1341-1347.
15. Simmonds P, Holmes EC, Cha TA, *et al.* Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J Gen Virol. 1993; 74(Pt 11): 2391-2399.
16. Cha TA, Kolberg J, Irvine B, *et al.* Use of a signature nucleotide sequence of hepatitis C virus for detection of viral RNA in human serum and plasma. J Clin Microbiol. 1992; 29(11): 2528-2534.
17. Tokita HS, Shrestha M, Okamoto H, *et al.* Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group. J Gen Virol. 1994; 75(Pt 4): 931-936.
18. Tokita HH, Okamoto H, Iizuka H, *et al.* Hepatitis C virus variants from Jakarta, Indonesia classifiable into novel genotypes in the second (2e and 2f), tenth (10a) and eleventh (11a) genetic groups. J Gen Virol. 1996; 77(Pt 2): 293-301.
19. Tokita HH, Okamoto H, Iizuka J, Kishimoto F, Tsuda Y, Mayumi M. The entire nucleotide sequences of three hepatitis C virus isolates in genetic groups 7-9 and comparison with those in the other eight genetic groups. J Gen Virol. 1998; 79(Pt 8): 1847-1857.
20. Kiprijanovska S. Prevalence and genotyping of HCV infection in patients belonging to different risk groups in the Republic of Macedonia. Prilozi. 2012; (submitted)
21. Simjanovska LJ, Porcu K, Amitov V, Efremov GD, Polenakovic M. Reverse transcriptase/polymerase chain reaction analysis of hemodialysis ultrafiltrates and sera of hepatitis C virus positive patients. Int J Artif Organs. 2004; 27(1): 35-39.
22. Popovski Z, Dimovski A, Simjanovska L, *et al.* Molecular detection and characterization of hepatitis C virus in the Republic of Macedonia. Makedonski medicinski pregled. 1996; 50(4):85-90.
23. Kiprijanovska S, Noveski P, Chalovska V, Polenakovic M, Plaseska-Karanfilska D. Association of *OASL* gene with chronic HCV therapy response to pegylated interferon-α and ribavirin. Proceedings of the 39th European Society for Artificial Organs Congress, Rostock, Germany, 26-29 September 2012; (submitted).

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

INTEGRATED GENOMIC ANALYSIS OF BREAST CANCERS

Addou-Klouche L[1,2], Adélaïde[1] J, Cornen S[1], Bekhouche I[1], Finetti P[1], Guille A[1], Sircoulomb F[1], Raynaud S[1], Bertucci F[1,3,4], Birnbaum D[1], Chaffanet M[1,*]

***Corresponding Author:** Max Chaffanet, Ph.D., HDR, Department of Molecular Oncology, Institut Paoli Calmettes, 232 Boulevard Sainte Marguerite, 13009 Marseille, France; Tel.: +33-(0)4-91-22-34-77; Fax: +33-(0)4-91-22-35-44; E-mail: CHAFFANETM@ipc.unicancer.fr

ABSTRACT

Breast cancer is the most frequent and the most deadly cancer in women in Western countries. Different classifications of disease (anatomoclinical, pathological, prognostic, genetic) are used for guiding the management of patients. Unfortunately, they fail to reflect the whole clinical heterogeneity of the disease. Consequently, molecularly distinct diseases are grouped in similar clinical classes, likely explaining the different clinical outcome between patients in a given class, and the fact that selection of the most appropriate diagnostic or therapeutic strategy for each patient is not done accurately. Today, treatment is efficient in only 70.0-75.0% of cases overall. Our repertoire of efficient drugs is limited but is being expanded with the discovery of new molecular targets for new drugs, based on the identification of candidate oncogenes and tumor suppressor genes (TSG) functionally relevant in disease. Development of new drugs makes therapeutical decisions even more demanding of reliable classifiers and prognostic/predictive tests. Breast cancer is a complex, heterogeneous disease at the molecular level. The combinatorial molecular origin and the heterogeneity of malignant cells, and the variability of the host background, create distinct subgroups of tumors endowed with different phenotypic features such as response to therapy and clinical outcome. Cellular and molecular analyses can identify new classes biologically and clinically relevant, as well as provide new clinically relevant markers and targets.

The various stages of mammary tumorigenesis are not clearly defined and the genetic and epigenetic events critical to the development and aggressiveness of breast cancer are not precisely known. Because the phenotype of tumors is dependent on many genes, a large-scale and integrated molecular characterization of the genetic and epigenetic alterations and gene expression deregulation should allow the identification of new molecular classes clinically relevant, as well as among the altered genes and/or pathways, the identification of more accurate molecular diagnostic, prognostic/predictive factors, and for some of them, after functional validation, the identification of new therapeutic targets.

Keywords: Breast cancers, Genome, Transcriptome, Epigenome, Oncogenes, Tumor suppressor genes

1 Marseille Cancer Research Center, Department of Molecular Oncology, UMR891 Inserm; Institut Paoli-Calmettes, Marseille, France
2 Biotoxicology Laboratory, Djillali Liabes University, Sidi-Bel-Abbès, Algeria
3 Department of Medical Oncology, Institut Paoli-Calmettes, Marseille, France
4 Université de la Méditerranée, Marseille, France

INTRODUCTION

High-Throughput Molecular Analyses in Breast Cancer and Translational Research. Unprecedented molecular characterization is possible using high-throughput molecular analyses, available at the DNA level with comparative genomic hybridization on microarrays (aCGH) [1-4], and at the RNA level, for expression profiling with DNA microarrays [5]. When these techniques emerged, expected applications were multiple in oncology, in both basic and translational research.

A number of studies have already shown the promising role of DNA microarray-based expression profiling in breast cancer translational research by identifying new clinically and biologically relevant intrinsic molecular subtypes (luminal A, luminal B, ERBB2+, basal, and normal-like) [6-7] and new prognostic and/or predictive gene signatures, whose predictive impact is superior to conventional histoclinical factors (for review, see [8]). Currently, three prognostic gene signatures are already commercially available: Oncotype DX (Genomic Health, Inc., Redwood City, CA, USA), MammaPrint (Agendia BV, Amsterdam, The Netherlands), and the HOXB13/ IL17BR (H/I) ratio (Theros H/ISM; bioTheranostics, San Diego, CA, USA). Others under development include the Intrinsic Gene Set, the Rotterdam Signature, the Wound Response Indicator, and the Invasive Gene Signature. Similarly, signatures predictive for response to specific therapies have been reported [9-12]. These prognostic or predictive signatures, once prospectively validated, will provide the opportunity to refine our therapeutic approach by individualizing treatment to patients'individual tumor profiles, likely contributing to significantly improve the clinical outcome (for review, see [13]).

The aCGH technology has been applied more recently to breast cancer. To date, some studies, including ours, have suggested a prognostic role of genomic data [14-16]. The integrative analysis of whole-genome expression and genomic data has revealed promising results for identifying candidate genes (identified as deregulated at the DNA and RNA levels simultaneously) associated with breast cancer or with specific features of disease [14,16-24].

For years, our laboratory has identified a large number of molecular alterations in recurrent breast cancer associated with: ***i)*** structural aberrations such as breakages [25-29], and ***ii)*** evaluated the clinical impact of the amplification [14,30,31]. We were among the first to demonstrate that the integrative analysis of whole-genome expression and genomic high resolution data are useful to identify new oncogenes and TSG specific to a clinical entity or a molecular subtype. Therefore, our comparative analyses of integrated profiles of breast cancers have been reported in basal and luminal tumors, two molecular subtypes of very different clinical courses [19], but also in particularly aggressive cancer: inflammatory breast cancer [32], breast cancers in young women (Raynaud *et al.,* in preparation), and *ERBB2* amplified breast cancers [33]. This laboratory was also one of the first to identify specific genomic markers of luminal B: *L3MBTL4* (18p11) [34] and *ZNF703* (8p12) [35] as potential TSG and oncogene, respectively.

Candidate Genes May Also be Transcriptionnally Deregulated Because of Epigenetic Alterations. The widespread deregulation of basic epigenetic profiles has emerged as a common phenotypic trait of cancer cells [36-38]. The epigenetic modifications include covalent tags added to nucleosome histone components [*e.g.,* acetylation of histone H3 and/or H4 (H3/4Ac) and/or various levels of methylation on lysine residues of histone H3 (H3K4/ K9me1/ 2/3), a non exhaustive list defined as the histone code], as well as methylation of CpG dinucleotides [39,40]. This applies particularly to CpG methylation profiles, whose modification has direct implication on many aspects of cell biology, namely cell division, survival, development and, consequently, oncogenesis. DNA methylation at regulatory regions of a gene, including promoter, generally leads to transcriptional silencing. CpG methylation-dependent silencing is now considered as an important mechanism of TSG inactivation in cancer cells, in addition to somatic genetic lesions [41]. DNA methylation changes in human cancers are complex and vary between different tumor types. Promoter methylation effectively represses transcription and occurs in many genes involved in human breast cancer development [42]. Among these, genes associated with cell cycle regulation (APC, RASSF1, RB, TFAP2A), or coding for steroid receptors (ESR1, PGR, RARα), suppressors (BRCA1, CDKN2A, CST6), and genes associated

with metastasis (CDH1, CEACAM6, PCDHGB6) and other genes such as *NRG1*. The majority of these affected genes are potential or known TSG [43]. Interestingly, there is also increasing evidence that methylation of regulatory regions of cancer-related genes can be one of the most prevalent molecular markers for human cancer diseases [44]. The potential clinical applications of DNA-methylation biomarkers may include diagnosis of neoplasm, tumor classification, prediction of response to treatment, or prognosis. DNA methylation status has thus been extensively studied in various molecular or clinical entities in breast cancers in order to better characterize them or improve their molecular classification [45-49].

In the continuity of our strategy, the high resolution DNA promoter methylation status will be analyzed on human promoter array (Agilent Technologies, Massy, France) and integrated to the genomic and gene expression data previously collected in the same set of 300 breast tumors. High-throughput molecular analyses of breast cancer have already revealed some part of their potential. Such integrated approaches could contribute to better understand the various levels of the dynamic molecular changes in the mammary oncogenesis and identify new markers.

REFERENCES

1. Solinas-Toldo S, Lampel S, Stilgenbauer S, *et al.* Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances.Genes Chromosomes Cancer. 1997; 20(4): 399-407.
2. Pinkel D, Segraves R, Sudar D, *et al.* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet. 1998; 20(2): 207-211.
3. van Beers EH, Nederlof PM. Array-CGH and breast cancer. Breast Cancer Res. 2006; 8(3): 210-219.
4. Curtis C, Lynch AG, Dunning MJ, *et al.* The pitfalls of platform comparison: DNA copy number array technologies assessed. BMC Genomics. 2009; 10: 588-610.
5. Bertucci F, Houlgatte R, Nguyen C, Viens P, Jordan BR, Birnbaum D. Gene expression profiling of cancer by use of DNA arrays: how far from the clinic? Lancet Oncol. 2001; 2(11): 674-682.
6. Perou CM, Sørlie T, Eisen MB, *et al.* Molecular portraits of human breast tumours. Nature. 2000; 406 (6797): 747-752.
7. Sørlie T, Perou CM, Tibshirani R, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001; 98(19): 10869-10874.
8. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. N Engl J Med. 2009; 360(8): 790-800.
9. Chang JC, Wooten EC, Tsimelzon A, *et al.* Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet. 2003; 362(9381):362-369.
10. Ayers M, Symmans WF, Stec J, *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol. 2004; 22(12): 2284-2293.
11. Bonnefoi H, Potti A, Delorenzi M, *et al.* Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncol. 2007; 8(12): 1071-1078.
12. Farmer P, Bonnefoi H, Anderle P, *et al.* A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009; 5(1): 68-74.
13. Dunn L, Demichele A. Genomic predictors of outcome and treatment response in breast cancer. Diagn Mol Ther. 2009; 13(2): 73-90.
14. Gelsi-Boyer V, Trouplin V, Adélaïde J, *et al.* Genome profiling of chronic myelomonocytic leukemia: frequent alterations of RAS and RUNX1 genes. BMC Cancer. 2008; 8: 299-313.
15. Bergamaschi A, Kim YH, Wang P, *et al.* Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer. 2006; 45(11): 1033-1040.
16. Chin K, de Vries S, Fridlyand J, *et al.* Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell. 2006; 10(6): 529-541.
17. Stransky N, Vallot C, Reyal F, *et al.* Regional copy number-independent deregulation of transcription in cancer. Nat Genet. 2006; 38(12): 1386-1396.
18. Neve RM, Chin K, Fridlyand J, *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006; 10(6): 515-527.
19. Adélaïde J, Finetti P, Bekhouche I, *et al.* Integrated profiling of basal and luminal breast cancers. Cancer Res. 2007; 67(24): 11565-11575.
20. Chin SF, Teschendorff AE, Marioni JC, *et al.* High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol. 2007; 8(10): R215-R231.

21. Bernard-Pierrot I, Gruel N, Stransky N, *et al.* Characterization of the recurrent 8p11-12 amplicon identifies PPAPDC1B, a phosphatase protein, as a new therapeutic target in breast cancer. Cancer Res. 2008; 68(17): 7165-7175.
22. Vincent-Salomon A, Lucchesi C, Gruel N, *et al.* Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. Clin Cancer Res. 2008; 14(7): 1956-1965.
23. André F, Job B, Dessen P, *et al.* Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. Clin Cancer Res. 2009; 15(2): 441-451.
24. Horlings HM, Lai C, Nuyten DS, *et al.* Integration of DNA copy number alterations and prognostic gene expression signatures in breast cancer patients. Clin Cancer Res. 2010; 16(2): 651-663.
25. Popovici C, Basset C, Bertucci F, *et al.* Reciprocal translocations in breast tumor cell lines: cloning of a t(3;20) that targets the FHIT gene. Genes Chromosomes Cancer. 2002; 35(3): 204-218.
26. Adélaïde J, Huang HE, Murati A, *et al.* A recurrent chromosome translocation breakpoint in breast and pancreatic cancer targets the heregulin/NRG1 gene at 8p12. Genes Chromosomes Cancer. 2003; 37(4): 333-345.
27. Huang HE, Chin S-F, Ginestier C, *et al.* A recurrent chromosome breakpoint in breast cancer at the NRG1/ neuregulin 1/heregulin gene. Cancer Res. 2004; 64(19): 6840-6844.
28. Letessier A, Ginestier C, Charafe-Jauffret E, *et al.* ETV6 gene rearrangements in invasive breast carcinoma. Genes Chromosomes Cancer. 2005; 44(1): 103-108.
29. Letessier A, Garrido-Urbani S, Ginestier C, *et al.* Correlated break at PARK2/FRA6E and loss of AF-6/ Afadin protein expression are associated with poor outcome in breast cancer. Oncogene. 2007; 26(2): 298-307.
30. Ginestier C, Bardou V-J, Popovici C, *et al.* Absence of FHIT protein expression correlates with adverse evolution in good prognosis localized breast cancer. Int J Cancer. 2003; 107(5): 854-862.
31. Letessier A, Sircoulomb F, Ginestier C, *et al.* Frequency, prognostic impact, and subtype association of 8p12, 8q24, 11q13, 12p13, 17q12, and 20q13 amplifications in breast cancers. BMC Cancer. 2006; 6(1): 245-257.
32. Bekhouche I, Finetti P, Adélaïde J, *et al.* Genome profiling of inflammatory breast cancer. PLoS One. 2011; 6(2): e16950-16962.
33. Sircoulomb F, Bekhouche I, Finetti P, *et al.* Genome profiling of ERBB2-amplified breast cancers. BMC Cancer. 2010; 10: 539-556.
34. Addou-Klouche L, Adélaïde J, Finetti P, *et al.* Loss, mutation and deregulation of L3MBTL4 in breast cancers. Mol Cancer. 2010; 9(1): 213-225.
35. Sircoulomb F, Nicolas N, Ferrari A, *et al. ZNF703* gene amplification at 8p12 specifies luminal B breast cancer. EMBO Mol Med. 2011; 3(3): 153-166.
36. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002; 3(6): 415-428.
37. Lund AH, van Lohuizen M. Epigenetics and cancer. Genes Dev. 2004; 18(19): 2315-2335.
38. Rice KL, Hormaeche I, Licht JD. Epigenetic regulation of normal and malignant hematopoiesis. Oncogene. 2007; 26(47): 6697-6714.
39. Goll MG, Bestor TH. Histone modification and replacement in chromatin activation. Genes Dev. 2002; 16(14): 1739-1742.
40. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci. 2006; 31(2): 89-97.
41. Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. Hum Mol Genet. 2007; 16(Spec 1): R50-R59.
42. Wildschwendter M, Jones PA. DNA methylation and breast carcinogenesis. Oncogene. 2002; 21(35): 5462-5482.
43. Mulero-Navarro S, Esteller M. Epigenetic biomarkers for human cancer: The time is now. Crit Rev Oncol Hematol. 2008; 68(1): 1-11.
44. Ordway JM, Budiman MA, Korshunova Y, *et al.* Identification of novel high-frequency DNA methylation changes in breast cancer. PLoS One. 2007; 2(12): e1314-1325.
45. Bediaga NG, Acha-Sagredo A, Guerra I, *et al.* DNA methylation epigenotypes in breast cancer molecular subtypes. Breast Cancer Res. 2010; 12(5): R77-R78.
46. Holm K, Hegardt C, Staaf J, *et al.* Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. Breast Cancer Res. 2010; 12 (3): R36-R51.
47. van der Auwera I, Yu W, Suo L, *et al.* Array-based DNA methylation profiling for breast cancer subtype discrimination. PLoS One. 2010; 5(9): e12616-12625.
48. Fackler MJ, Umbricht CB, Williams D, *et al.* Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. Cancer Res. 2011; 71(19): 6195-6207.
49. Hill VK, Ricketts C, Bieche I, *et al.* Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity. Cancer Res. 2011; 71(8): 2988-2999.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

MOLECULAR GENETICS OF BREAST AND OVARIAN CANCER: RECENT ADVANCES AND CLINICAL IMPLICATIONS

Bogdanova N[1,2], Dörk T[2,*]

***Corresponding Author:** Thilo Dörk, Ph.D., Hannover Medical School, Gynaecology Research Unit (OE 6411), Carl-Neuberg-Str. 1, D-30625 Hannover, Germany; Tel.: +49-511-532-6075; Fax: +49-511-532-6081; E-mail: doerk.thilo@mh-hannover.de

ABSTRACT

Over the last few years, evidence has been accumulated that several susceptibility genes exist that differentially impact on the lifetime risk for breast or ovarian cancer. High-to-moderate penetrance alleles have been identified in genes involved in DNA double-strand break signaling and repair, and many low-penetrance susceptibility loci have been identified through genome-wide association studies. In this review, we briefly summarize present knowledge about breast and ovarian cancer susceptibility genes and discuss their implications for risk prediction and therapy.

Keywords: Breast carcinoma, Ovarian carcinoma, Germ-line mutations, Chromosomal instability

REVIEW

Epidemiological and genetic linkage studies of mul-tiple-case families have guided the identification of *BRCA1* and *BRCA2* as the first genes in which mutations strongly predispose to breast and ovarian cancer. However, hereditary breast and ovarian cancer (HBOC) syndrome only represents the extreme end of a wide spectrum of genetically influenced breast or ovarian carcinomas. During the last few years, evidence has been accumulated that several susceptibility genes exist [1-3]. Their mutations have differential impact according to the minor allele frequencies and the magnitude of the allelic effect. We briefly summarize our present knowledge about breast and ovarian cancer susceptibility genes and discuss their implications for risk prediction and therapy.

Rare Mutations With a High-to-Moderate Penetrance. *BRCA1* and *BRCA2*: The prototypic *BRCA1* and *BRCA2* mutations confer a very high lifetime risk for breast cancer in the range of 55.0-85.0% for *BRCA1* and 35.0-60.0% for *BRCA2*, compared with an about 10.0% general population risk. Mutations are usually truncating, although pathogenic missense mutations have also been described. Lifetime risk for ovarian cancer is also high and may be up to 40.0% for *BRCA1* mutation carriers. There seems to be allele-specific expressivity as some of the mutations appear to confer higher risks for ovarian cancer than others [4-6]. Both the risks for breast and ovarian cancer can also be modified by additional gene loci such as single nucleotide polymorphisms (SNPs) in *RAD51* or *BNC2* [7,8] (see below).

The *PALB2* Mutation. Subsequently, *PALB2* has been identified as a breast cancer susceptibility gene [9, 10]. The PALB2 protein bridges BRCA1

[1] Clinic of Radiation Oncology, Hannover Medical School, Hannover, Germany
[2] Clinic of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany

and BRCA2 and synergizes in their function in recombinational DNA repair. Mutations in *PALB2* predispose to breast cancer and gastric cancer, and the penetrance for breast cancer in Finnish multiple-case families has been found to be as high as for *BRCA2* mutations [11]. There is less evidence that *PALB2* mutations may also constitute ovarian cancer susceptibility alleles. Although *PALB2* founder mutations have been identified in ovarian cancer patients from Poland and Russia, they are still rare in these populations [12,13].

The *RAD51 Paralogs*. Downstream of BRCA1, BRCA2 and PALB2, the RAD51 protein is a key mediator of homologous recombination, and a regulatory variant c.–98G>C (also known as135G>C) in *RAD51* acts as a genetic modifier of *BRCA2* mutations [7]. Mutation analyses in further genes of RAD51 paralogs have subsequently uncovered *RAD51C* and *RAD51D* as susceptibility genes for breast and ovarian cancer [14-16]. The initial data indicate that these mutations are specifically associated with a family history of ovarian cancer. However, mutations in *RAD51C* and *RAD51D* are collectively very rare and their penetrance and tumor spectrum remains to be fully explored.

Additional Fanconi Anemia Genes. Since some breast and ovarian cancer susceptibility alleles, *e.g.,* in *BRCA2*, *PALB2* or *RAD51C*, cause Fanconi anemia (FA) in the homozygous state, it seemed reasonable to assess further FA genes for their role in breast and ovarian cancer. So far, mutations of the *BRIP1* gene have been associated with breast cancer [17], although their precise risk estimates await further studies. There is no evidence implicating the FA core proteins in breast cancer, suggesting that proteins at the heart of the downstream homologous recombination machinery are the major factors for breast and ovarian cancer susceptibility.

Familial Lobular Breast Cancer. Familial lobular breast cancer has been associated with rare germ-line mutations in *CDH1*, the gene for E-cadherin [18,19]. While mutations in *CDH1* are also causative for diffuse gastric cancer, there is no evidence to implicate *CDH1* in ovarian cancer yet.

Rare Syndromes Including Breast Cancer. Some rare syndromes include the occurrence of breast cancer as part of the disease spectrum. These include Li-Fraumeni (*TP53*), Muir-Torre Syndrome (*MSH2*), Cowden's Disease (*PTEN*), Peutz-Jeghers Syndrome (*LKB1*), and Ataxia-telangiectasia (*ATM*, see below). Although these syndromes are generally rare, they need to be kept in mind if a breast cancer patient presents with a more complex disorder or family history.

Ataxia-Telangiectasia. Twenty-five years ago, it was shown that blood relatives of patients with the neurode-generative disorder ataxia-telangiectasia (A-T) face an increased breast cancer risk [20]. The gene mutated in ataxia-telangiectasia, *ATM*, encodes a master protein kinase that orchestrates the cellular response to DNA double-strand breaks and controls, *via* phosphorylation, hundreds of proteins involved in cell cycle control, repair and apoptosis, among them BRCA1, BRCA2, TP53, CHEK2 and many other tumor suppressors [21]. Truncating mutations in *ATM* appear to confer an about 3-fold increased breast cancer risk to heterozygous carriers [22-24]. While the homozygous condition of A-T is rare, heterozygotes may account for 0.5-1.0% of the population.

MRE11A/RAD50/NBN. Chromosome breaks are sensed and the ATM protein is activated *via* the MRN complex consisting of the proteins MRE11A, RAD50, and NBN. The *NBN* gene underlies Nijmegen Breakage Syndrome (NBS), which is most prevalent in Eastern Europe due to a Slavic founder mutation. While biallelic mutations cause NBS, heterozygous carriers face an about 3-5 fold increased breast cancer risk [25-27]. Similarly, biallelic mutations in *RAD50* give rise to a NBS-like disorder, whereas heterozygotes for a Finnish founder mutation are predisposed towards breast cancer [28]. *MRE11A* also is a gene for an A-T-like disorder, although there has been only one study to associate *MRE11A* mutations with breast cancer so far [29]. None of the three genes have been extensively investigated in ovarian cancer, but germ-line mutations in any of them were identified in a recent large sequencing study [30].

CHEK2. Checkpoint kinase 2 is a major target of ATM and itself phosphorylates further tumor suppressor proteins, including p53 and BRCA1, in response to DNA damage [31]. *CHEK2* had initially been found mutated in Li-Fraumeni patients and one of these mutations, 1100 delC, has subsequently been associated with familial breast cancer [32,33]. Heterozygous carriers for 1100delC have been reported with a 2- to 3-fold increase in breast cancer

risk [34], with rare homozygotes being found at much higher risk [35]. In Eastern Europe, two further truncating mutations IVS-II-1 (G>A) and *CHEK2*dele9,10 (5 kb) have been associated with at least similarly high breast cancer risks, whereas the missense mutation p.I157T has a lower penetrance [25,36,37]. There has not been conclusive evidence for a strong association of *CHEK2* mutations with ovarian cancer, but their association with additional malignancies suggests a more general role in cancer predisposition [38].

Polymorphic Variants With Low Penetrance. Several polymorphic loci are known today that influence the risk of breast and/or ovarian cancer. This has been achieved through genome-wide association studies (GWAS) of SNPs by large consortia during the past 5 years. The published GWAS efforts have so far uncovered over 20 genomic loci for breast cancer [39-45] and six loci for serous epithelial ovarian cancer [46-48] at a genome-wide significance level. All these loci harbor low-penetrance alleles with allelic odds ratios of less than 1.5. As these loci still explain only a small part of the heritable fraction and further large-scale studies are presently underway, it is likely that these numbers will increase very rapidly. Most of the hitherto identified loci appear to be specific for either breast or ovarian carcinomas. For example, the gene for fibroblast growth factor receptor 2, *FGFR2* [39], harbors variants associated with breast but not ovarian cancer, and the gene for basonuclin-2, *BNC2* [46], harbors variants associated with ovarian but not breast cancer. Nevertheless, there is a minor group of shared loci that appear to influence both breast and ovarian cancer risk. Such genes include *BABAM1* (encoding a BRCA1 binding partner also known as MERIT40), *TERT* (encoding a component of telomerase), and the protooncogene *MYC* on chromosome 8q24. Interestingly, a closer inspection of the 8q24 locus indicated that the associations with either breast or ovarian cancer were caused by independent variants at the same locus which may be explained by tissue-specific regulation of gene expression [49]. As a caveat, a GWAS roughly localizes but does not identify the causal variant, and in several cases, there is more than one candidate gene in the region spanned by the associated linkage disequilibrium block or under putative regulatory control of the identified locus. Fine-mapping approaches in different ethnic populations and gene expression studies are presently being used to further trace down the true predisposing gene variants. Copy number variants have also been investigated in a GWAS but this did not detect a significant association for breast cancer [50].

Implications for Risk Prediction and Therapy. The identification of mutations in individuals from families with HBOC makes it possible to predict the age-dependent risk for different cancers, including recurrence risks in the already affected, and to appropriately counsel the patient and her blood relatives. This may lead to an increased surveillance or targeted prevention including medication (such as tamoxifen) or preventive surgery (such as prophylactic oophorectomy). In many countries, this is available to patients with an over 35.0% lifetime risk such as *BRCA1* or *BRCA2*, and possibly *PALB2* mutation carriers, whereas a more restrained position is taken for patients with moderate-penetrance mutations conferring an about 3-fold increase in breast cancer risk such as *ATM* or *CHEK2*. Although it might also be suggested that for these mutations, the female carriers should benefit from increased surveillance, large studies on the efficacy of such measures are still lacking. No further consequences are considered for patients carrying common risk alleles at polymorphic loci, as these risks are too small individually to be clinically meaningful. With the identification of many more low-risk loci, however, it may become possible to calculate combinatorial risks that could be useful in a stratified approach for cancer prevention in the future [51,52].

Identifying the genetic basis of breast or ovarian cancer in the individual patient might have further prognostic and therapeutic implications. For a long time, breast cancer therapy has been guided by the presence or absence of gene products such as hormone receptors or HER2/neu. Such gene expression profiles are partly determined by germ-line mutations such as *BRCA1* mutations which are frequently associated with triple-negative breast cancers [53], but breast cancer pathology also seems to be influenced by low-penetrance variants such as in *FGFR2*, which is strongly associated with estrogen-receptor positive disease [54]. Studies are presently underway to investigate whether SNP

profiling could thus be of prognostic value, and there are new drugs being developed that target additional breast cancer pathways such as those mediated by fibroblast growth factor receptors [55]. Recent reports further indicate that the outcome of ovarian cancer therapy is significantly influenced by the *BRCA1/BRCA2* mutational status. In these multi-center studies, mutation carriers showed an improved survival, probably due to a higher benefit from the usually applied platinum-based therapy that activates a DNA repair pathway defective in BRCA1 or BRCA2 deficient tumors [56]. Another recent approach to improve targeted therapy is based on the concept of "synthetic lethality" as exemplified by the introduction of PARP1 inhibitors into breast and ovarian cancer treatment of patients with *BRCA1* or *BRCA2* mutations, and probably beyond [57]. Here, the idea is to inhibit a repair pathway that can still be compensated by backup pathways in normal but not tumor cells. The apparent success of this concept has stimulated the targeting of other repair pathways in parallel to those known to be defective in breast or ovarian carcinomas. For example, inhibition of ATR, which is a backup kinase of ATM, has been reported to be particularly effective in tumors with *TP53* or *ATM* mutations [58]. Though promising, such substances still need further development until they can be tested in clinical practice. It is the hope that with many more genes identified, a deeper understanding of breast and ovarian cancer development and progression, together with the ability of gene-based stratification, will ultimately lead to an improved and individually tailored therapy for the benefit of each patient.

REFERENCES

1. Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet. 2008; 9: 321-345.
2. Ghoussaini M, Pharoah PD. Polygenic susceptibility to breast cancer: current state-of-the-art. Future Oncol. 2009; 5(5): 689-701.
3. Gayther SA, Pharoah PD. The inherited genetics of ovarian and endometrial cancer. Curr Opin Genet Dev. 2010; 20(3): 231-238.
4. Risch HA, McLaughlin JR, Cole DE, *et al.* Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. Am J Hum Genet. 2001; 68(3): 700-710.
5. Thompson D, Easton D, Breast Cancer Linkage Consortium. Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. Am J Hum Genet. 2001; 68(2): 410-419.
6. Bogdanova NV, Antonenkova NN, Rogov YI, Karstens JH, Hillemanns P, Dörk T. High frequency and allele-specific differences of BRCA1 founder mutations in breast cancer and ovarian cancer patients from Belarus. Clin Genet. 2010; 78(4): 364-372.
7. Antoniou AC, Sinilnikova OM, Simard J, *et al.* RAD51 135G>C modifies breast cancer risk among BRCA2 mutation carriers: results from a combined analysis of 19 studies. Am J Hum Genet. 2007; 81(6): 1186-1200.
8. Ramus SJ, Kartsonaki C, Gayther SA, Pharoah PD, Sinilnikova OM, Beesley J. Genetic variation at 9p22.2 and ovarian cancer risk for BRCA1 and BRCA2 mutation carriers. J Natl Cancer Inst. 2011; 103(2): 105-116.
9. Rahman N, Seal S, Thompson D, *et al.* PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet. 2007; 39(2): 165-167.
10. Erkko H, Xia B, Nikkilä J, *et al.* A recurrent mutation in PALB2 in Finnish cancer families. Nature. 2007; 446(7133): 316-319.
11. Erkko H, Dowty JG, Nikkilä J, *et al.* Penetrance analysis of the PALB2 c.1592delT founder mutation. Clin Cancer Res 2008;14(14):4667-4671.
12. Dansonka-Mieszkowska A, Kluska A, Moes J, *et al.* A novel germline PALB2 deletion in Polish breast and ovarian cancer patients. BMC Med Genet. 2010; 11: 20.
13. Prokofyeva D, Bogdanova N, Bermisheva M, *et al.* Rare occurrence of PALB2 mutations in ovarian cancer patients from the Volga-Ural region. Clin Genet. 2012; 82(1): 100-101.
14. Meindl A, Hellebrand H, Wiek C, *et al.* Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet. 2010; 42(5): 410-414.
15. Pelttari LM, Heikkinen T, Thompson D, *et al.* RAD51C is a susceptibility gene for ovarian cancer. Hum Mol Genet. 2011; 20(16): 3278-3288.
16. Loveday C, Turnbull C, Ramsay E, Hughes D, Ruark E, Frankum JR. Germline mutations in RAD51D confer susceptibility to ovarian cancer. Nat Genet. 2011; 43(9): 879-882.
17. Seal S, Thompson D, Renwick A, *et al.* Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nat Genet. 2006; 38(11): 1239-1241.

18. Masciari S, Larsson N, Senz J, *et al.* Germline E-cadherin mutations in familial lobular breast cancer. J Med Genet. 2007; 44(11): 726-731.
19. Xie ZM, Li LS, Laquet C, *et al.* Germline mutations of the E-cadherin gene in families with inherited invasive lobular breast carcinoma but no diffuse gastric cancer. Cancer. 2011; 117(14): 3112-3117.
20. Swift M, Reitnauer PJ, Morrell D, Chase CL. Breast and other cancers in families with ataxia-telangiectasia. N Engl J Med. 1987; 316(21): 1289-1294.
21. Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer. 2003; 3(3): 155-168.
22. Renwick A, Thompson D, Seal S, *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet. 2006; 38(8): 873-875.
23. Pylkäs K, Tommiska J, Syrjäkoski K, *et al.* Evaluation of the role of Finnish ataxia-telangiectasia mutations in hereditary predisposition to breast cancer. Carcinogenesis. 2007; 28(5): 1040-1045.
24. Bogdanova N, Cybulski C, Bermisheva M, *et al.* A nonsense mutation (E1978X) in the ATM gene is associated with breast cancer. Breast Cancer Res Treat. 2009; 118(1): 207-211.
25. Górski B, Cybulski C, Huzarski T, *et al.* Breast cancer predisposing alleles in Poland. Breast Cancer Res Treat. 2005; 92(1): 19-24.
26. Steffen J, Nowakowska D, Niwinska A, *et al.* Germline mutations 657del5 of the NBS1 gene contribute significantly to the incidence of breast cancer in Central Poland. Int J Cancer. 2006; 119(2): 472-475.
27. Bogdanova N, Feshchenko S, Schürmann P, *et al.* Nijmegen Breakage Syndrome mutations and risk of breast cancer. Int J Cancer. 2008; 122(4): 802-806.
28. Heikkinen K, Rapakko K, Karppinen SM, *et al.* RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis. 2006; 27(8): 1593-1599.
29. Bartkova J, Tommiska J, Oplustilova L, *et al.* Aberrations of the MRE11-RAD50-NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. Mol Oncol. 2008; 2(4): 296-316.
30. Walsh T, Casadei S, Lee MK, *et al.* Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci USA. 2011; 108(44): 18032-18037.
31. Nevanlinna H, Bartek J. The CHEK2 gene and inherited breast cancer susceptibility. Oncogene. 2006; 25(43): 5912-5919.
32. Meijers-Heijboer H, van den Ouweland A, Klijn J, *et al;* CHEK2-Breast Cancer Consortium. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet. 2002; 31(1): 55-59.
33. Vahteristo P, Bartkova J, Eerola H, *et al.* A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. Am J Hum Genet. 2002; 71(2): 432-438.
34. CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. Am J Hum Genet. 2004; 74(6): 1175-1182.
35. Adank MA, Jonker MA, Kluijt I, *et al.* CHEK2* 1100delC homozygosity is associated with a high breast cancer risk in women. J Med Genet. 2011; 48(12): 860-863.
36. Cybulski C, Wokołorczyk D, Huzarski T, *et al.* A deletion in CHEK2 of 5,395 bp predisposes to breast cancer in Poland. Breast Cancer Res Treat. 2007; 102(1): 119-122.
37. Bogdanova N, Enssen-Dubrowinskaja N, Feshchenko S, *et al.* Association of two mutations in the CHEK2 gene with breast cancer. Int J Cancer. 2005; 116(2): 263-266.
38. Cybulski C, Górski B, Huzarski T, *et al.* CHEK2 is a multiorgan cancer susceptibility gene. Am J Hum Genet. 2004; 75(6): 1131-1135.
39. Easton DF, Pooley KA, Dunning AM, *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007; 447(7148): 1087-1093.
40. Ahmed S, Thomas G, Ghoussaini M, *et al.* Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet. 2009; 41(5): 585-590.
41. Antoniou AC, Wang X, Fredericksen ZS, *et al.* A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet 2010;42(10):885-892.
42. Turnbull C, Ahmed S, Morrison J, *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet. 2010; 42(6): 504-507.
43. Haiman CA, Chen GK, Vachon CM, *et al.* A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet. 2011; 43(12): 1210-1214.
44. Long J, Cai Q, Sung H, *et al.* Genome-wide association study in East Asians identifies novel susceptibility loci for breast cancer. PLoS Genet. 2012; 8(2): e1002532.
45. Ghoussaini M, Fletcher O, Michailidou K, *et al.* Genome-wide association analysis identifies three

new breast cancer susceptibility loci. Nat Genet. 2012; 44(3): 312-318.
46. Song H, Ramus SJ, Tyrer J, *et al.* A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet. 2009; 41(9): 996-1000.
47. Goode EL, Chenevix-Trench G, Song H, *et al.* A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet. 2010; 42(10): 874-879.
48. Bolton KL, Tyrer J, Song H, *et al.* Common variants at 19p13 are associated with susceptibility to ovarian cancer. Nat Genet. 2010; 42(10): 880-884.
49. Ghoussaini M, Song H, Koessler T, *et al.* Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst. 2008; 100(13): 962-966.
50. Wellcome Trust Case Control Consortium. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature. 2010; 464(7289): 713-720.
51. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med. 2008; 358(26): 2796-2803.
52. Pashayan N, Duffy SW, Chowdhury S, *et al.* Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. Br J Cancer. 2011; 104 (10): 1656-1663.
53. Robertson L, Hanson H, Seal S, *et al.* BRCA1 testing should be offered to individuals with triple-negative breast cancer diagnosed below 50 years. Br J Cancer. 2012; 106(6): 1234-1238.
54. Broeks A, Schmidt MK, Sherman ME, *et al.* Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium. Hum Mol Genet. 2011; 20(16): 3289-3303.
55. Hynes NE, Dey JH. Potential for targeting the fibroblast growth factor receptors in breast cancer. Cancer Res. 2010; 70(13): 5199-5202.
56. Bolton KL, Chenevix-Trench G, Goh C, *et al.* Association between BRCA1 and BRCA2 mutations and survival in women with invasive epithelial ovarian cancer. JAMA. 2012; 307(4): 382-390.
57. Dedes KJ, Wilkerson PM, Wetterskog D, Weigelt B, Ashworth A, Reis-Filho JS. Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations. Cell Cycle. 2011; 10(8): 1192-1199.
58. Reaper PM, Griffiths MR, Long JM, *et al.* Selective killing of ATM- or p53-deficient cancer cells through inhibition of ATR. Nat Chem Biol. 2011; 7(7): 428-430.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

GENETIC VARIATION OF THE *BRCA1* AND *BRCA2* GENES IN MACEDONIAN PATIENTS

Maleva I[1], Madjunkova S[1], Bozhinovski G[1], Smickova E[2], Kondov G[3], Spiroski Z[3], Arsovski A[4], Plaseska-Karanfilska D[1,*]

***Corresponding author:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

ABSTRACT

The most significant and well characterized genetic risk factors for breast and/or ovarian cancer are germline mutations in the *BRCA1* and *BRCA2* genes. The *BRCA1* and *BRCA2* gene mutations strikingly increase breast cancer risk, suggesting that polymorphisms in these genes are logical candidates in seeking to identify low penetrance susceptibility alleles. The aim of this study was to initiate a screen for *BRCA1/2* gene mutations in order to identify the genetic variants in the Republic of Macedonia, and to evaluate the association of several single nucleotide polymorphisms (SNPs) in these genes with breast cancer risk. In this study, we included 100 patients with invasive breast cancer from the Republic of Macedonia, classified according to their family history and 100 controls. The methodology included direct sequencing, single nucleotide primer extension method and multiplex ligation probe amplification (MLPA) analysis, all followed by capillary electrophoresis (CE) on an ABI PRISM™ 3130 Genetic Analyzer. We identified a total of seven carriers of mutations in the *BRCA1/2* genes. None of the tested polymorphisms was associated with sporadic breast cancer risk, however, polymorphism rs8176267 in *BRCA1* and N372H in *BRCA2* showed an association with breast cancer risk in patients with at least one family member with breast cancer.

Keywords: *BRCA1* and *BRCA2* genes, Breast cancer, Macedonian patients, Polymorphisms.

INTRODUCTION

The most significant and well characterized genetic risk factors for breast and/or ovarian cancer are germline mutations in the *BRCA1* (17q chromosome) [1] and *BRCA2* (13q chromosome) [2] genes. Other relevant genes, such as *CHEK2, NBS1, PALB2, BRIP1*, *etc.,* also contribute to hereditary breast cancer, although their impact appears to be more population-specific [3]. It has been estimated that 5.0-10.0% of all breast cancer and 10.0-15.0% of ovarian cancer patients carry mutations on one of the *BRCA* genes [4]. The prevalence of the *BRCA1/2* gene mutation carriers in the general population is approximately 0.2% (1/500), however, it can vary significantly in different countries and ethnic groups due to founder effects [5]. The mutations in these high-penetrance genes confer a high lifetime risk of breast and ovarian cancer. Women with an inherited

[1] Research Center for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Science and Arts, Skopje, Republic of Macedonia
[2] Clinic for Oncology and Radiotherapy, Skopje, Republic of Macedonia
[3] Clinic for Thoracic Surgery, Skopje, Republic of Macedonia
[4] Re-Medika General Hospital, Skopje, Republic of Macedonia

BRCA1 gene mutation have a 65.0-80.0% risk of developing breast cancer and 37.0-62.0% of developing ovarian cancer over their lifetime, while *BRCA2* gene mutation carriers have a 45.0-85.0% risk for breast cancer and 11.0-23.0% for ovarian cancer [6]. The identification of *BRCA1* and *BRCA2* gene mutation carriers is therefore a critical step in individualized risk assessment [7]. Once a mutation is identified in a given family, a very informative predictive (or presymptomatic) oncogenetic test can be offered to all adult family members. Moreover, oncogenetic testing is becoming the powerful therapeutic predictive tool, as new targeted therapeutic opportunities, such as poly(ADP-ribose) (PARP) inhibitors emerge [8] and chemosensitivity to platinum-based therapy is constantly reported [9]. It is now evident that in the near future the demand for rapid *BRCA1/2* gene mutation testing will increase. However, a full *BRCA1* and *BRCA2* gene screening still remains a labor- and time-consuming challenge due to the large size of the genes and the high diversity of mutations and variants of unknown significance. On the other hand, the distribution of known *BRCA1* and *BRCA2* gene mutations is well documented worldwide. Several recent reviews have summarized the evidence that the *BRCA1/2* gene mutation spectrum in given countries and ethnic communities is limited to a few founder mutations [4,5,10]. To date, no systematic study has assessed the distribution of *BRCA1/2* gene mutations in the Macedonian population. We aimed to initiate screening for *BRCA1/2* gene mutations in order to identify the genetic variants common in the Republic of Macedonia.

The fact that *BRCA1* and *BRCA2* gene mutations drastically increase breast cancer risk suggests that polymorphisms in these genes could represent low penetrance susceptibility alleles [11]. Whether common polymorphisms contribute to disease risk has not yet been thoroughly evaluated. The importance of these common variants is still conflicting and more data on large cohorts are needed to better understand their significance. We present data on several single nucleotide polymorphisms (SNPs) including allele frequencies and association with breast cancer risk.

MATERIALS AND METHODS

We included 100 patients with invasive breast cancer from the Republic of Macedonia in this study. The patients were referred to us from the Clinic for Oncology, Skopje and the Re-Medika General Hospital, Skopje, Republic of Macedonia. Patients were classified into three main groups, according to their family history: ***group 1)*** patients with two or more close relatives with breast cancer (n = 19); ***group 2)*** patients with only one affected relative with breast cancer (n = 31); and ***group 3)*** patients with no family history (sporadic cases) (n = 50). The control group consisted of healthy women from the general population (n = 100).

The DNA was isolated from peripheral EDTA blood samples using standard proteinase K/SDS digestion followed by phenol chloroform extraction. All patients were screened for six mutations in the *BRCA1* gene (185delAG, C61G, E368X, 4154delA, 4184del4 and 5382insC) and four in the *BRCA2* gene (D2723G, 3034del4, 5950delCT and 9326insA) by a single nucleotide primer extension assay utilizing the ABI PRISM™ SNaPshot Multiplex Kit (Life Technologies, Carlsbad, CA, USA) following the manufacturer's instructions (manuscript in preparation). Patients from the first group and patients younger than 40 years from the second group (n = 30) were screened for mutations in all coding sequences of the *BRCA1* and *BRCA2* genes by direct sequencing using the ABI PRISM™ Big Dye Terminator v.1.1 Kit (Life Technologies). Multiplex ligation probe amplification (MLPA) was used for the detection of large rearrangements in these genes using commercially available kits from MRC Holland, Amsterdam, The Netherlands. For the case-control association study of seven common variants in *BRCA1* [rs1799949 (S694S), rs799917 (P871L), rs16941 (E1038G), rs16942 (E1138G), rs8176267, rs8176166 and rs3737559) and one in *BRCA2* (rs144848 (H372N)] with breast cancer risk, we also used single nucleotide primer extension.

RESULTS AND DISCUSSION

Point mutations in the *BRCA* genes are the most common deleterious mutations in familial breast cancer patients. Complete sequencing remains the gold standard for initial mutation identification. However, large rearrangements in these genes have been described in a significant proportion of breast cancer families and are responsible for up to one-third of the identifiable BRCA mutations in a cer-

Table 1. Mutations in the *BRCA1* and *BRCA2* genes detected by sequencing analysis.

Gene	BIC[a] MutationName	HGVS Nomenclature	Exon	Amino Acid Change	Location in Gene	Gender
BRCA1	**C61G**	c.181T>G	5	Cys→Gly	RING domain	F
BRCA1	**E368X**	c.1102G>T	11	Glu→Stop	RING domain	F
BRCA1	**del ex5-8**		5-8		ex5,6-RING domain	F
BRCA1	**5382insC**	c.5266dupC	20	Glu→Stop	BRCT domain	F
BRCA1	**del ex23**		23		BRCT domain	F
BRCA2	**D2723G**	c.8167G>C	18	Asp→His	DNA binding domain	M

[a] BIC: Breast Cancer Information Core.

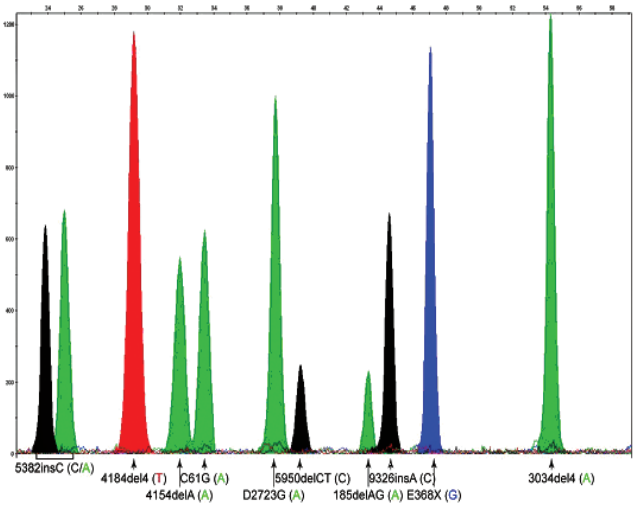


Figure 1. Single nucleotide primer extension assay. Analysis of a patient carrying the 5382insC mutation in the *BRCA1* gene: the fluorescent peaks formed by specific primer extension products are labeled below the electropherogram; labels correspond to mutation names (see Table 2).

Table 2. Specific primer extension products data for most common mutations in the *BRCA* genes.

No.	Mutation	Nucleotide Change	SNaPshot Result (N/M)	SNaPshot Fragment Size N (bp)	SNaPshot Fragment Size M (bp)
1	**5382insC**	A/C	A/C	25	24
2	**4184del4**	T/G	T/G	29	n.d.
3	**4154delA**	A/G	A/G	32	n.d.
4	**C61G**	T/G	T/G	33.5	32
5	**D2723G**	A/G	A/G	38	37
6	**5950delC**	C/A	C/A	39	n.d.
7	**185delAG**	A/T	A/T	43	n.d.
8	**9326insA**	C/A	C/A	44.5	n.d.
9	**E368X**	G/T	G/T	47	48
10	**3034del4**	A/G	A/G	54	n.d.

n.d.: not done.

tain population [12]. In our group of familial breast cancer patients (group 1 as defined in Materials and Methods), we identified a total of seven carriers of mutations: four point mutations and two large deletions in the *BRCA1* gene and a point mutation in the *BRCA2* gene (Table 1). Mutations appear to be evenly distributed across the coding sequence of the genes. Bearing in mind that certain mutations have been observed to be common to specific populations, we designed an assay for detection of the most common mutations in the Slavic populations. Our aim was to expand the mutation screen to breast cancer patients regardless of their family history. To this end, we developed a single nucleotide primer extension as a rapid and economical one-tube test for genetic testing of hereditary breast cancer that can be applied to a wider population setting (Figure 1, Table 2). We screened all sporadic patients and did not identify any mutations until now. More analyses including direct sequencing are needed in order to assess the distribution of mutations in the Macedonian population. This is important because it will allow the development of effective mutation-specific tests for the common mutations in the future. In patients with a strong familial history of breast cancer (n = 6), we performed mutational screening in all coding exons of the *PALB2* gene using the high resolution melting (HRM) method. These analyses were performed at the Gynecology Unit, Hannover

Table 3. Specific primer extension products data for most analyzed SNPs in the *BRCA* genes.

No.	Mutation	Nucleotide Change	SNaPshot Result (N/M)	SNaPshot Fragment Size N (bp)	SNaPshot Fragment Size M (bp)
1	**rs16942**	A/G	T/C	27	25
2	**rs8176267**	A/G	A/G	30	28
3	**rs16941**	A/G	T/C	31	28
4	**rs3737559**	A/G	A/G	31	33
5	**rs799917**	C/T	C/T	34	35.5
6	**rs8176166**	A/G	T/C	39	38
7	**rs1799949**	C/T	C/T	44	45
8	**rs144848**	T/G	A/C	50	49

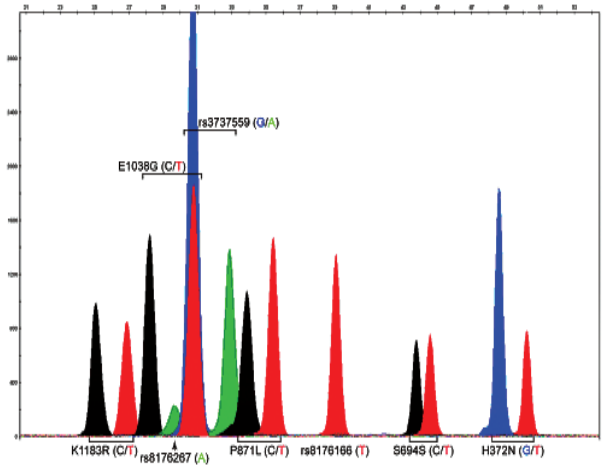


Figure 2. Single nucleotide primer extension assay for detection of selected SNPs in the *BRCA* genes: the fluorescent peaks formed by specific primer extension products are labeled below the electropherogram; labels correspond to SNP names (see Table 3).

Medical School, Hannover, Germany. We found three already published polymorphisms and one potentially damaging variant.

Many studies are focused on rare, highly penetrant germline mutations in *BRCA* genes that strongly predispose women to a familial form of breast cancer. However, there is a possibility that common germline variation in coding and non coding regions may also contribute to predisposition to breast cancer. In the present study, we tested seven common variations in the *BRCA* genes (Figure 2, Table 3) in all our patients in comparison to the controls. Our results showed that none of the polymorphisms tested were associated with the risk of sporadic breast cancer (group 3) suggesting that the variations *per se* do not play a significant role in the development of sporadic breast cancer. However, polymorphism rs8176267 in the *BRCA1* gene showed an association with breast cancer risk when we analyzed the results for patients with at least one family member with breast cancer (groups 1 and 2 combined) *vs.* controls [$p = 0.0151$; OR (odds ratio) (95% CI) (95% confidence interval) = 2.31 (1.16-4.61)]. These results are in concordance with published data [13]. Recent meta-analysis suggests that the *BRCA2* N372H allele may be a low-penetrant risk factor for developing breast cancer [14], however, there is conflicting evidence regarding the role of this variant as a modifier of breast cancer risk. We observed that N372H is associated with slightly increased risk in patients with a family member with breast cancer [$p = 0.0081$; OR (95%CI) =2.37 (1.24-4.56)] [14].

Further analyses on larger cohorts of patients and controls are needed in order to build a high-quality database of genetic *BRCA1/2* gene variants in the Macedonian population, and to obtain accurate estimates as to the association of various polymorphisms with breast cancer risk.

ACKNOWLEDGMENTS

We are grateful to Professor Thilo Dörk for the opportunity of performing mutational screening in the *PALB2* gene in Macedonian breast cancer patients with a strong familial history of the disease (negative for *BRCA1/2* gene mutations) during the stay of Ms. Ivana Maleva at the Gynecology Unit, Hannover Medical School, Hannover, Germany. Ms. Maleva's 4-week stay was supported by the FP7 project No. 229458 from the European Commission.

REFERENCES

1. Miki Y, Swensen J, Shattuck-Eidens D, *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994; 266(5182): 66-71.
2. Wooster R, Bignell G, Lancaster J, *et al.* Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995; 378(6559): 789-792.
3. Walsh T, King MC. Ten genes for inherited breast cancer. Cancer Cell. 2007; 11(2): 103-5.
4. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat Rev Cancer. 2007; 7(12): 937-948.
5. Ferla R, Calo V, Cascio S, *et al.* Founder mutations in BRCA1 and BRCA2 genes. Ann Oncol. 2007; 18 (Suppl 6): vi93-98.
6. Balmana J, Diez O, Castiglione M. BRCA in breast cancer: ESMO clinical recommendations. Ann Oncol. 2009; 20(Suppl 4): 19-20.
7. De Greve J, Sermijn E, De Brakeleer S, Ren Z, Teugels E. Hereditary breast cancer: from bench to bedside. Curr Opin Oncol. 2008; 20(6): 605-613.
8. Rouleau M, Patel A, Hendzel MJ, Kaufmann SH, Poirier GG. PARP inhibition: PARP1 and beyond. Nat Rev Cancer. 2010; 10(4): 293-301.
9. Byrski T, Gronwald J, Huzarski T, *et al.* Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. J Clin Oncol. 2010; 28(3): 375-379.
10. Neuhausen SL. Founder populations and their uses for breast cancer genetics. Breast Cancer Res. 2000; 2(2): 77-81.
11. Freedman ML, Penney KL, Stram DO, *et al.* A haplotype-based case-control study of BRCA1 and sporadic breast cancer risk. Cancer Res. 2005; 65(16): 7516-7522.
12. Ewald IP, Ribeiro PL, Palmero EI, Cossio SL, Giugliani R, Ashton-Prolla P. Genomic rearrangements in BRCA1 and BRCA2: A literature review. Genet Mol Biol. 2009; 32(3): 437-446.
13. Cox DG, Kraft P, Hankinson SE, Hunter DJ. Haplotype analysis of common variants in the BRCA1 gene and risk of sporadic breast cancer. Breast Cancer Res. 2005; 7(2): R171-R175.
14. Qiu LX, Yao L, Xue K, *et al.* BRCA2 N372H polymorphism and breast cancer susceptibility: a meta-analysis involving 44,903 subjects. Breast Cancer Res Treat. 2010; 123(2): 487-490.

Proceedings of the MACPROGEN Final Conference held
at Ohrid, Republic of Macedonia, March 29-April 1 2012

MicroRNAs IN BREAST CANCER —OUR INITIAL RESULTS

Popovska-Jankovic K[1], Noveski P[1], Chakalova L[1], Petrusevska G[2], Kubelka K[3], Plaseska-Karanfilska D[1]

***Corresponding Author:** Professor Dr. Dijana Plaseska-Karanfilska, Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Krste Misirkov 2, Skopje 1000, Republic of Macedonia; Tel: +389(0)2 3235410; Fax: +389 (0)2 3115434; E-mail: dijana@manu.edu.mk

ABSTRACT

MicroRNAs (miRNAs) are small [~21 nucleotide (nt)] non coding RNAs (ncRNAs) that regulate gene expression posttranscriptionally. About 3.0% of human genes encode for miRNAs, and up to 30.0% of human protein coding genes may be regulated by miRNAs. Currently, more than 2000 unique human mature microRNAs are known. MicroRNAs play a key role in diverse biological processes including development, cell proliferation, differentiation and apoptosis. These processes are commonly dysregulated in cancer, implicating miRNAs in carcinogenesis, where they act as tumor supressors or oncogenes. Several miRNAs are associated with breast cancer. Here we present our initial results of miRNA analyses of breast cancer tissues using quantitative real time-polymerase chain reaction (ReTi-PCR) (qPCR) involving stem-loop reverse transcriptase (RT) primers combined with TaqMan® PCR and miRNA microarray analysis.

Keywords: Breast cancer, microRNA (miRNA), Microarray, Real time-polymerase chain reaction (ReTi-PCR), Stem-loop reverse transcriptase (RT) primers.

[1] Research Center for Genetic Engineering and Biotechnology "Georgi D. Efremov," Macedonian Academy of Sciences and Arts, Skopje, Republic of Macedonia
[2] Institute of Pathology, Faculty of Medicine, University "Stt. Cyril and Methodius," Skopje, Republic of Macedonia
[3] Clinical Hospital "Acibadem Sistina," Skopje, Republic of Macedonia

INTRODUCTION

MicroRNAs (miRNAs) are small [~21-nucleotide (nt)] non coding RNAs (ncRNAs) that mediate posttranscriptional gene regulation by pairing with the 3' untranslated region (3'UTR) of messenger RNAs (mRNAs), acting as translational repressors, and regulating gene expression posttranscriptionally. After the discovery of the first mi-RNA in the roundworm *Caenorhabditis elegans*, these short regulatory RNAs have been found to be an abundant class of RNAs in plants, animals and DNA viruses. About 3.0% of human genes encode for miRNAs, and up to 30.0% of human protein coding genes may be regulated by miRNAs [1]. Currently more than 2000 unique human mature microRNAs are known [2]. MicroRNAs play a key role in diverse biological processes including development, cell proliferation, differentiation and apoptosis [1]. Thus, potentially all cellular pathways may be governed by miRNAs, which may contribute to the fine tuning of gene expression on a global level. The importance of miRNAs in gene regulation will be better appreciated when their function or deregulation, or that of the cellular machinery mediating their biosynthesis and function, will be identified among the underlying causes of several genetic disorders. Indeed, it is easy to conceive that protein over expression resulting from defective miRNA-based mRNA regulation may compromise normal cell function and cause

genetic diseases [3]. Accordingly, altered miRNA expression is likely to contribute to human disease, including cancer [1]. In cancer, miRNAs function as regulatory molecules, acting as oncogenes or tumor suppressors [4]. Amplification or over expression of miRNAs can down regulate tumor suppressors or other genes involved in cell differentiation, thereby contributing to tumor formation by stimulating proliferation, angiogenesis and invasion, *i.e.,* they act as oncogenes. Similarly, miRNAs can down regulate different proteins with oncogenic activity, *i.e.,* they act as tumor suppressors [5,6].

Several miRNA are associated with breast cancer. It has been shown that there are differences not just between normal and breast cancer tissue, but also between different breast cancer subtypes [7-12].

Several methods for global miRNA profiling are currently in common use. These include quantitative real time-polymerase chain reaction (ReTi-PCR) (qPCR) involving stem-loop reverse transcriptase (RT) primers combined with TaqMan® PCR (Life Technologies, Carlsbad, CA, USA) analysis, qPCR with locked nucleic acid primers (Exiqon, Vedback, Denmark), qPCR using poly(A) tailing (Qiagen, Hilden, Germany; Stratagene, La Jolla, CA, USA), high-throughput sequencing of small RNA libraries and microarray analysis. We have recently initiated a study of microRNAs in breast cancer tissues with a main aim to search for breast cancer diagnostic and prognostic markers.

MATERIALS AND METHODS

Fresh frozen tissues (normal and malignant) from patients with breast cancer were obtained from the Institute of Pathology, Medical Faculty, Skopje, Republic of Macedonia. Histopathological data were obtained from all patients. The protocol used for miRNA quantification in tissue samples included three separate procedures: extraction of total RNA, RT- and ReTi-PCR assays. RNA extraction was performed using an RNeasy Kit (Qiagen). Quality and quantity of total RNAs were checked on 1.0% agarose gels and a nanodrop spectrophotometer. RNA samples were dissolved in RNase-free water and stored at –80°C. MicroRNA quantitation was performed by stem-loop RT-PCR followed by Taq-Man® PCR analysis [13] using TaqMan® MicroRNA Reverse Transcription Kit, TaqMan® Universal PCR Master Mix and five TaqMan® MicroRNA Assays (miR-155, miR21, miR-125b and miR-145 and RNU6b as a control gene) (Life Technologies). The RT-PCR mix was made on ice in a final volume of 15 μL, following the manufacturer's protocol. Thermal cycling conditions were: 30 min. at 16°C, 30 min. at 42°C and 5 min. at 85°C. The ReTi-PCR assay was performed in duplicates in a total volume of 20 μL consisting of 7.7 μL ddH_2O, 1.0 μL 20 × TaqMan® Small RNA Assay, 10.0 μL 2 × TaqMan® Universal PCR Master Mix and 1.3 μL RT reaction product (Life Technologies). Thermal cycling conditions were: enzyme activation and initial denaturation at 95°C for 10 min., followed by 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C. The expression of each miRNA relative to RNU6b was determined using the ΔΔCt method. MicroRNA microarray analysis was done using a complete labeling and hybridization kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer's protocol [14].

RESULTS AND DISCUSSION

Here we present our initial results of miRNA analysis of breast cancer and normal tissues that included quantitative ReTi-PCR (qPCR) involving stem-loop RT primers combined with TaqMan® PCR (Life Technologies) and miRNA microarray analyses. A total of 35 patients with breast cancer were analyzed using ReTi-PCR of four miRNAs (miR-155, miR-21, miR-125b and miR-145). Our initial results showed that miR-155 and miR-21 are up regulated and miR-125b and miR-145 are down regulated in breast cancer. Our results are in agreement with other published studies for miRNA expression profiling in breast cancer [7,8,11,12,15,16]. MicroRNA microarray analysis was performed in several patients and the data were analyzed by R and MeV statistical programs. Nine miRNAs were differentialy expressed, of which seven were up regulated (miR-155, miR-146a, miR-150, miR-210, miR-21, miR-106b and miR-142-3p) and two were down regulated (miR-139-5p and miR-320c). These up regulated miRNAs were also shown to be differentially expressed in breast cancer by other authors [7-9,15]. Nevertheless, additional testing and analyses are needed to establish accurate and precise miRNA markers that will contribute to the improvement of diagnosis and prognosis in breast cancer patients.

ACKNOWLEDGMENTS

We are grateful to Drs. Max Chaffanet and Pascal Finetti for the MeV statistical microarray analyses that were performed during the stay of Ms. Katerina Popovska-Jankovic at the Marseilles Cancer Research Center, Marseille, France. Ms. Popovska-Jankovic's 3-week stay was supported by the FP7 project No. 229458 from the European Commission.

REFERENCES

1. Sassen S, Miska EA, Caldas C. MicroRNA-implications for cancer. Virchows Arch. 2008; 452(1): 1-10.
2. miRBase: the microRNA database (http://microrna. sanger. ac.uk).
3. Ouellet DL, Perron, MP, Gobeil LA, Plante P, Provost P. MicroRNAs in gene regulation: when the smallest governs it all. J Biomed Biotechnol. 2006; 2006(4): 69616.
4. Kent OA, Mendell JT. A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes. Oncogene 2006; 25(46): 6188-6196.
5. Hammond SM. MicroRNAs as oncogenes. Curr Opin Genet Dev. 2006; 16(1): 4-9.
6. Shenouda SK, Alahari SK. MicroRNA function in cancer: oncogene or a tumor suppressor? Cancer Metastasis Rev. 2009; 28(3-4): 369-378.
7. Iorio MV, Ferracin M, Liu CG, *et al.* MicroRNA gene expression deregulation in human breast cancer. Cancer Res. 2005; 65(16): 7065-7070.
8. Mattie MD, Benz CC, Bowser J, *et al.* Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. Mol Cancer. 2006; 5: 24.
9. Blenkiron C, Goldstein LD, Thorne NP, *et al.* MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. Genome Biol. 2007; 8(10): R214.
10. Yan LX, Huang XF, Shao Q, *et al.* MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. RNA. 2008; 14(11): 2348-2360.
11. Hui ABY, Shi W, Boutros PC, *et al.* Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. Lab Invest. 2009; 89(5): 597-606.
12. Lowery AJ, Miller N, Devaney A, *et al.* MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/*neu* receptor status in breast cancer. Breast Cancer Res. 2009; 11(3): R27.
13. Chen C, Ridzon DA, Broomer AJ, *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res. 2005; 33(20): e179.
14. miRNA Microarray System with miRNA Complete Labeling and Hyb Kit http://www.genomics.agilent. com/files/Manual/G4170-90011_miRNA_ Complete_ 2.2.pdf.
15. Volinia S, Calin GA, Liu CG, *et al.* A microRNA expression signature of human solid tumors defines gene target. Proc Natl Acad Sci USA. 2006; 103(7): 2257-2261.
16. Sempere LF, Christensen M, Silahtaroglu A, *et al.* Altered microRNA expression confined to specific epithelial cell subpopulations in breast cancer. Cancer Res. 2007; 67(24): 11612-11620.

INFORMATION FOR AUTHORS

1. Aims and scope of the journal

Balkan Journal of Medical Genetics (BJMG) was founded in 1998, and encompasses all aspects of human genetics: cytogenetics, molecular genetics, clinical genetics, immunogenetics, oncogenetics, pharmacogenetics, population genetics, genetic screening and diagnosis of monogenic and polygenic diseases, prenatal and preimplantation genetic diagnosis, genetic counseling, advances in treatment and prevention.

The journal is supported by an internationa Editorial Board.

The BJMG welcomes the submission of manuscripts that meets the general criteria of significance and scientific excellence, and will publish original articles, case studies, reviews, surveys, opinions, commentaries and essays.

2. Identification Statement

BJMG is published to the web as articles become available. The journal is published twice a year in June and November by Macedonian Academy of Sciences and Arts.

3. Subscription Price

Print subscription price: 120 US$ for public institutions and 50 US$ for individuals.

Reprints are available. The price is 50 US$ for 50 reprints and 100 US$ for 100 reprints.

Payments can be made by cheque and bank transfer to the following bank account.

Bank Name and Address: City Bank New York, 111 Wall Street, New York, NY 10043

Bank Account Number: (NBRM MK 2X) 36086854

Bank Account Owner: National Bank of the Republic of Macedonia; Macedonian Academy of Sciences and Arts - ICGIB, Skopje

Swift Code: CITI USD 33

The transfer costs (taxes and commissions) are at the expense of the receiving bank and are included in the above mentioned price. Please send copy of the bank certificate to confirm the payment.

4. Submission of manuscripts:

Manuscripts written in English should be submitted online on www.bjmg.edu.mk or sent by e-mail as an attached file to the Editorial Office:

Dijana Plaseska-Karanfilska (Editor)
Macedonian Academy of Sciences and Arts
Av. Krste Misirkov 2, POB 428
1000 Skopje, Republic of Macedonia
E-mail: manu.icgib@gmail.mk

The text files should be prepared using MS Word or Rich Text Format (RTF) and should be clearly labeled with the first author's names. Each manuscript must be accompanied by a statement that it has not been published or submitted for publication elsewhere. Authors are responsible for obtaining permission to reproduce copyrighted material from other sources and are required to sign an agreement for the transfer of copyright to the publisher. All accepted manuscripts, artwork, and photographs become the property of the publisher.

5. Manuscript Review

Manuscripts are evaluated by two or more expert reviewers. Authors are encouraged to suggest the names and contacts (including e-mail address) of prospective reviewers. Authors are notified of the disposition of a manuscript six weeks after receipt.

7. Preparation of manuscript

- All manuscripts must be accompanied by a cover letter from the corresponding author stating that the manuscript represents original, unpublished material not under editorial consideration elsewhere, and that ethical guidelines were followed in the conduct of the research.

- All manuscripts must be written in English (the Vancouver style) and the text typed double spaced on one side of the paper only with wide margins of at least 2.5 cm on all four sides. Uniform requirements for manuscripts may be found in Br Med J 1991; 302:338-341, and N Engl J Med 1991; 324: 424-428.

- The title page should contain: title of article, abbreviated running title of no more than 30 characters, author's name, department, institution, city and country, name and full postal address, telephone, fax and e-mail of the corresponding author. All pages including the title page should be numbered.

- The abstract of no more than 250 words should start on page 2. A list of 3-6 key words or short phrases should be given in alphabetic order below the abstract.

- The main text should be arranged in the following components: introduction, materials and methods, results, discussion, acknowledgments, references, figure legends and tables.

- Cite each reference, figure and table in numerical order - ex.[1-5]; (Table 1); (Fig.1).

- References (not exceed 30-40 in number) should be numbered in order in which they are first mentioned in the text. The surnames of all authors followed by initials should be given. Abbreviate journal names in accordance with the Index Medicus system (Br Med J 1988;296:401-405).

Examples: Du Manoir S, Speicher M, Joos S, Schroc E, Popp S, Dohner H, Kovacs G, Robert-Nicaud M, Lichter P, Cremer T; Detection of complete and partial chromosome gain and losses by comparative genomic in situ hybridization. Hum Genet 1993; 90:590-610.

- Place all the tables at the end of your file. Start each table on new page, complete with title and footnotes (if any). Number the tables (Arabic numerals) in the order in which they are cited in the text. Always separate the individual columns using tabulators, not using the space bar.

- Figures should be prepared on separate pages. Number the figures in Arabic numerals in the order in which they are cited in the text.