

## Meeting report series

### **Report of the Rare Genetic Diseases: Diagnosis and Discovery Workshop *Partnership Opportunities with Central/Eastern Europe and the Middle East***

Prague, Czech Republic  
December 3, 2013

#### **Organization**

Organized by: Kym Boycott (Department of Paediatrics, University of Ottawa and Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada) and Milan Macek

Hosted by: Milan Macek (Department of Biology and Medical Genetics, Charles University Prague-2<sup>nd</sup> Faculty of Medicine and University Hospital Motol)

#### **Speakers**

Dr Nicoleta Andreescu, Timisoara, Rumania

Dr David Atlan, Montreux, Switzerland

Dr Jana Behuňová, Vienna, Austria (also representing Slovakia)

Dr Kym Boycott, Ottawa, Canada, co-chair

Dr Michael Brudno, Toronto, Canada

Prof Xavier Estivill, Barcelona, Spain

Dr Sergey Kutsev, Moscow, Russia

Prof Milan Macek, Prague, Czech Republic, co-chair

Dr Halyna Makukh, Lviv, Ukraine

Dr Béla Méleg, Pécs, Hungary

Dr Dragica Radojkovič, Belgrade, Serbia

Dr Anna Rajab, Muscat, Sultanate of Oman

Dr Jadranka Sertič, Zagreb, Croatia

Dr Ewa Zietkiewicz, Poznan, Poland

Dr. Maria Tzetis and Dr. Jan Traeger Synodinos, Athens, Greece (via Skype link)

## Agenda

1. Welcome and introduction (Kym Boycott, Michael Brudno, Milan Macek)
2. IRDiRC: overview (Kym Boycott)
3. Canada's efforts in rare diseases (Kym Boycott)
4. Geuvadis Consortium (Xavier Estivill)
5. Presentations from Central/Eastern European and Middle Eastern partners
  - a. Victor Babes University of Medicine and Pharmacy, Department of Genetics, Timisoara, Rumania (Nicoleta Andreescu)
  - b. Centre of Pathobiochemistry and Genetics, Institute of Medical Genetics, Vienna, Austria (Jana Behuňová)
  - c. Department of Paediatric Surgery, Children's University Hospital and Faculty of Medicine, Comenius University, Bratislava, Slovakia (Jana Behuňová)
  - d. Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia (Sergey Kutsev)
  - e. Institute of Hereditary Pathology, Academy of Medical Sciences, Lviv, Ukraine (Halyna Makukh)
  - f. Department of Medical Genetics, University of Pécs, Pécs, Hungary (Béla Mélégh)
  - g. Department of Medical Genetics, Medical School, University of Athens, Athens, Greece (Jan Traeger Synodinos)
  - h. Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia (Dragica Radojkovič)
  - i. The National Genetics Centre, Muscat, Sultanate of Oman (Anna Rajab)
  - j. Clinical Institute of Laboratory Diagnosis, Zagreb University School of Medicine and Clinical Hospital Centre, Zagreb, Croatia (Jadranka Sertič)
  - k. Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poland (Ewa Zietkiewicz)
  - l. National Coordination Centre for Rare Diseases, Charles University Prague, University Hospital Motol, Prague, Czech Republic (Milan Macek)
6. Mechanisms to enable international data sharing
  - a. Cafe Variome (David Atlan)
  - b. PhenoTips (Michael Brudno)
  - c. PhenomeCentral (Michael Brudno)
  - d. MatchMaker (Michael Brudno)
7. Next steps



## REPORT

### 1. Welcome and introduction

Kym Boycott, from the University of Ottawa, Michael Brudno, from the University of Toronto, and Milan Macek, from the Charles University Prague and University Hospital Motol of Prague, welcomed all the participants to the Rare Genetic Diseases: Diagnosis and Discovery Workshop. This workshop aims to create partnership opportunities between Central/Eastern Europe, the Middle East, and Canada.

### 2. IRDiRC overview

Kym Boycott presented the International Rare Diseases Research Consortium (IRDiRC, <http://www.irdirc.org/>). The IRDiRC teams up researchers and organizations investing in rare disease research in order to achieve two main objectives by the year 2020, namely to deliver 200 new therapies for rare diseases and means to diagnose most rare diseases.

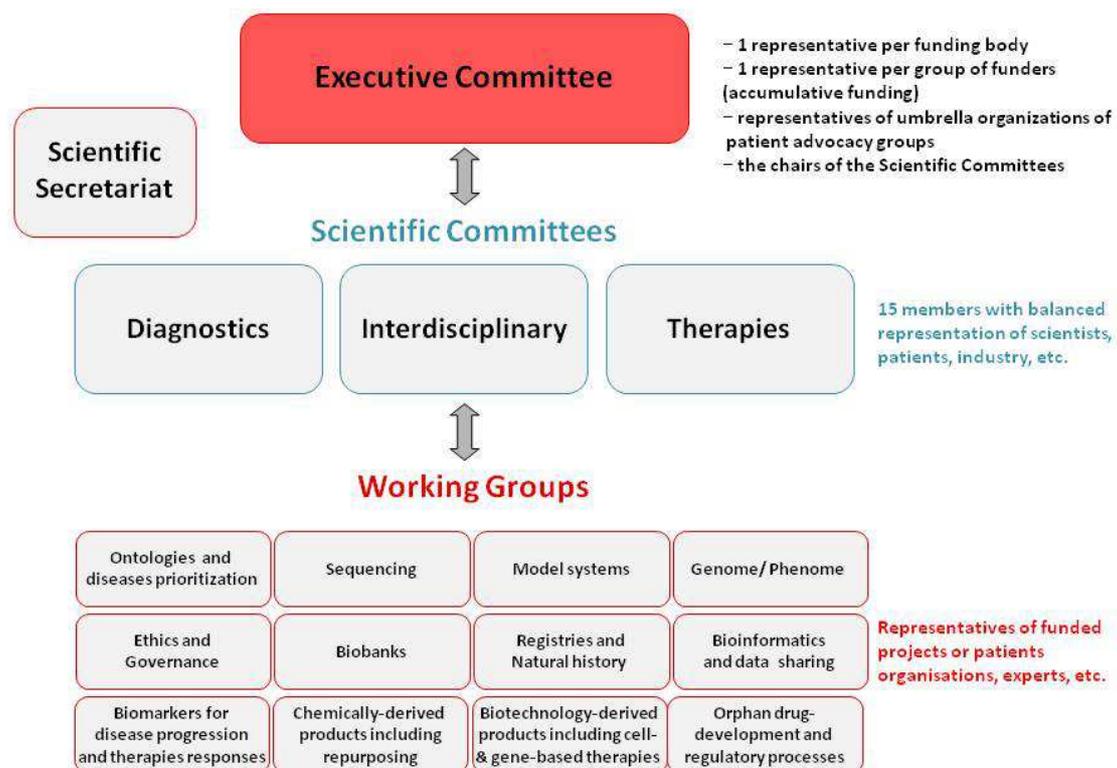
A number of grand challenges are being addressed through collaborative actions to reach these 2020 goals such as:

- ▶ Establishing and providing access to harmonized data and samples
- ▶ Performing the molecular and clinical characterization of rare diseases
- ▶ Boosting translational, preclinical and clinical research
- ▶ Streamlining ethical and regulatory procedures

The IRDiRC represents a large global effort of international collaboration between different countries. In April 2011, IRDiRC was officially established and launched during the second preparatory workshop in Bethesda (USA). IRDiRC is composed of 37 funding agencies coming from 13 different countries. The IRDiRC Scientific Secretariat facilitates the IRDiRC goals by providing organizational support for the implementation of the IRDiRC roadmap in close collaboration with the different committees and all the stakeholders.

IRDiRC is governed by an Executive Committee, three Scientific Committees (the Diagnostics Scientific Committee, the Interdisciplinary Scientific Committee, and the Therapies Scientific Committee), and twelve working groups. The Executive Committee is composed of one representative per funding body or funding group, representatives of patient umbrella organizations, and the Chair of each of the three Scientific Committees. Each Scientific Committee is composed of approximately 15 members with a balanced expertise and representation from academia, patient organizations, diagnostics, pharmaceutical industry, and regulatory bodies. They advise the Executive Committee on research priorities and progress made from a scientific point of view. The working groups are composed of representatives from funded projects that contribute to IRDiRC objectives within the scientific domain of the working group. They cooperate to ensure synergies between all research projects, by exchanging results, expertise, experiences and information.

Summary of the IRDiRC structure is below.



IRDiRC intends to monitor progress towards its goals through a set of direct and indirect indicators. For example, IRDiRC tries to collect all the research projects focused on rare diseases. Another example is that IRDiRC also tries to count the number of new genes identified as being at the origin of rare diseases. High-level cooperation from different countries will be necessary to achieve IRDiRC goals.

### 3. Canada's efforts in rare diseases

CARE4RARE (<http://care4rare.ca/>) is a pan-Canadian collaborative team of geneticists, clinicians, bioinformaticians and researchers building upon the infrastructure and discoveries of the FORGE Canada project (April 2011-March 2013, Genome Canada and Canadian Institutes of Health Research). The main goal of CARE for RARE is to improve diagnostics and clinical care for patients (and families) suffering from these diseases. The funding envelope for FORGE and CARE4RARE is >15 million Canadian dollars.

The FORGE project is complete and the presentation focused on this project. FORGE had three main activities:

- ▶ Identification of rare diseases present in the Canadian population
- ▶ Use exome sequencing to identify novel rare disease genes
- ▶ Centralization of data within a single national platform

>350 rare diseases were proposed for study, and 258 were selected. The diseases selected were:

- ▶ Present in childhood or adolescence
- ▶ Likely monogenic with unknown gene
- ▶ Present in at least one Canadian patient

Three strategies have been implemented for gene discovery:

- ▶ Multiple alleles strategy: unrelated patients or unrelated families with the same disorder
- ▶ Mapping data strategy: consanguineous families or autosomal dominant families with multiple affected members
- ▶ Compound heterozygous strategy: autosomal recessive disorders, affected sibpairs

Of the 258 disorders studied using exome sequencing, 149 disorders were solved and 109 disorders are still unsolved. Of the 258 disorders studied:

- ▶ 1/3 of the genes were novel genes
- ▶ 1/3 of the were known disease genes: includes atypical phenotypes, conflation of two diseases, or new mechanisms
- ▶ 1/3 of the disorders need more work to be solved

#### **4. Geuvadis Consortium**

GEUVADIS (Genetic European Variation in Health and Disease, <http://www.geuvadis.org>) is a European medical sequencing Consortium coordinated by Xavier Estivill at the Centre for Genomic Regulation in Barcelona. It is funded for a period of 36 months by the European Commission under its 7<sup>th</sup> Framework Programme. Investigators at 17 institutions across Europe and the US have joined together in an effort to establish standards for human genome sequencing diagnostic practice and data/variant analysis.

GEUVADIS has two main objectives:

- ▶ Set up the standards for RNA sequencing (establish protocols, share data)
- ▶ Optimise exome sequencing and create a dedicated pan-European exome variant server

GEUVADIS aims to:

- ▶ Establish standards for quality control and sequence data assessment
- ▶ Develop models for data storage, exchange, and access
- ▶ Standardize biological and medical sequence data interpretation, handling, and analysis
- ▶ Establish ethical standards related to sequence variation-based phenotype prediction
- ▶ Develop effective dissemination and training forums and protocols

GEUVADIS also aims to develop efficient forums to exchange medical information as well as any technological insights its members generate in pursuit of their specific objectives.

The RNA-sequencing work package of the GEUVADIS project has combined transcriptome and genome sequencing data by performing mRNA and small RNA sequencing on 465 lymphoblastoid cell line (LCL) samples from 5 populations of the 1000 Genomes Project: the CEPH (CEU), Finns (FIN), British (GBR), Toscani (TSI) and Yoruba (YRI). Of these samples, 423 were part of the 1000 Genomes Phase 1 dataset with low-coverage whole genome and high-coverage exome sequencing data, and the remaining 42 are part of the later phases of 1000 Genomes with Omni 2.5M SNP array data available at the time of this study.

The GEUVADIS RNA-sequencing data are freely and openly available. The main portal for accessing the data is EBI ArrayExpress (accessions E-GEUV-1, E-GEUV-2, E-GEUV-3). For visualisation of the results, the GEUVADIS Data Browser ([www.ebi.ac.uk/Tools/geuvadis-das](http://www.ebi.ac.uk/Tools/geuvadis-das)) was created where quantifications and QTLs can be viewed, searched, and downloaded.

## **5. Short summary of presentations from Central/Eastern European and Middle Eastern partners**

### **a. Victor Babes University of Medicine and Pharmacy, Department of Genetics, Timisoara, Rumania**

This department collaborates with rare disease patient associations in Rumania. It also collaborates internationally with the Karolinska Institute in Stockholm, Sweden, and with the Oslo University Hospital in Norway. It has an infrastructure for conventional and molecular genetics analyses. Some of the cases studied by the department have been published in international journals and have been incorporated into international databases.

Below are some examples of patients that the Department has studied:

- ▶ Prader-Willi syndrome
- ▶ Neurofibromatosis
- ▶ Duchenne muscular dystrophy
- ▶ Rare cancers
- ▶ Osteosarcomas
- ▶ Leukemias

This department is working hard to become a centre of excellence for rare diseases in the country. Recently, it has drafted a proposal to obtain funds for units dedicated to proteomics, bioinformatics, biobanks, and translational research.

### **b. Centre of Pathobiochemistry and Genetics, Institute of Medical Genetics, Vienna, Austria**

This Centre is part of the Medical University of Vienna, in Austria. Its director is Dr. M. Hengstschläger, and the director of clinical genetics is Prof. F. Laccone.

Its clinicians and research focus on:

- ▶ Monogenic diseases
- ▶ Hereditary cancers
- ▶ Rett syndrome

- ▶ Marfan syndrome

Five medical doctors and one postgraduate student work at this Centre. The Vienna Centre performs genetic counselling and clinical syndromology.

Some of the molecular techniques available at the Centre comprise:

- ▶ Cytogenetics/molecular cytogenetics/FISH
- ▶ Mutation analysis (by PCR RFLP, Sanger/NGS sequencing, MLPA, arrayCGH, etc.)

Patients' spectra comprise:

- ▶ Unknown syndromes
- ▶ Predictive genetic tests (genetic cancers, Huntington's disease)
- ▶ Infertility
- ▶ Prenatal diagnostics

**c. Department of Paediatric Surgery, Children's University Hospital and Faculty of Medicine, Comenius University, Bratislava, Slovakia**

Slovakia has a population of over five million and approximately 55,000 births/year. Approximately 10% of the population is composed by the Roma population, which has a higher birth rate and high rate of consanguinity, contributing to the increased prevalence of rare diseases. Clinical diagnosis of rare diseases is difficult in the Roma population since their contact with doctors is very rare. Thus far, the only genetic laboratory that diagnoses rare diseases is in the capital, Bratislava, which means that patients have to travel from all parts of the country. The technological capacity is rather limited, and doesn't include mitochondrial analysis or microarray testing.

**d. Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia**

The Research Centre for Medical Genetics of the Russian Academy of Medical Sciences was founded in 1968 on the basis of the Institute of Medical Genetics of the Soviet Academy of Medical Sciences, which was established in response to the activity of an earlier generation of Russian human geneticists. The Centre is headed by Prof Evgeni K. Ginter. It is composed of twelve laboratories and two Departments: the Federal Centre of Medical Genetics, and the Cystic Fibrosis Federal Centre for diagnosis and treatment of cystic fibrosis. The Centre's laboratories perform the following analyses:

- ▶ Molecular biology including biotechnology group, and metabolic hereditary diseases
- ▶ DNA diagnostics (around 4,200 diseases diagnosed; development of DNA testing protocols)
- ▶ Molecular genetics of complex inherited diseases
- ▶ Genetic epidemiology
- ▶ Environmental genetics
- ▶ Genetics of reproduction disorders
- ▶ Epigenetics
- ▶ Prenatal diagnosis
- ▶ Mutagenesis

- ▶ Genetics of stem cells
- ▶ Constitutional cytogenetics
- ▶ Human population genetics of the Russian Federation

The Centre has a biobank of approximately 5,000 samples of DNA, blood, fibroblasts, plasma, and urine from around 15,000 donors and comprising for instance:

- ▶ Cystic fibrosis
- ▶ Deafness
- ▶ Inherited disorders
- ▶ Duchenne muscular dystrophy
- ▶ Metabolic disorders
- ▶ Lysosomal storage disorders

Unfortunately, there is not enough money to proceed to NGS-based diagnostics, thus cooperation with different laboratories (domestic and foreign) is needed.

#### **e. Institute of Hereditary Pathology, Academy of Medical Sciences, Lviv, Ukraine**

There are approximately 48 million inhabitants in Ukraine, with 10 million inhabitants living in the Western part of the country serviced by the Lviv institute. In the entire country, 24 Medical Genetic clinics within 24 “oblasts” (districts or primary administrative units) and 6 Medical Genetic centres are providing services to rare disease patients and their families. The Institute (created in 1998) has 48 staff members and 34 students in training. It performs prenatal and clinical diagnostics for chromosomal and single genes disorders. In addition, there is a biobank with more than 3,000 samples including multiple rare disorders (e.g 285 patients with cystic fibrosis). Examined rare diseases comprise for example:

- ▶ Cystic fibrosis
- ▶ Nijmegen breakage syndrome (common in Slavic populations)
- ▶ Spinal muscular atrophy
- ▶ Ataxia telangiectasia
- ▶ Duchenne/Becker muscular dystrophy
- ▶ Dwarfism

The Institute has expertise in clinical syndromology, but needs to acquire facilities for Sanger and/or NGS sequencing, MLPA, arrayCGH or FISH. The Institute would also need additional specialists to cope with interpretation and increasing demand for laboratory tests.

#### **f. Department of Medical Genetics, University of Pécs, Pécs, Hungary**

This department is a diagnostic, treatment and research facility specialized in hereditary and acquired rare disorders. It offers advanced diagnostics, prevention and treatment for such diseases. The department has an up to date laboratory that provides cytogenetic, molecular genetics and molecular cytogenetic testing. The laboratory not only accepts samples of local, but also of international origin. The genetic counselling facility, in addition to clinical syndromology, is specialized in both common and rare diseases.

The department runs a national registry for congenital disorders (6,000 – 7,000 cases per year) and has a biobank (since 1990), in which the number of samples collected from patients suffering from rare genetic disorders is now reaching over 10,000 cases. It is a member of the European BBMRI.eu consortium. The department serves as the Hungarian national centre for rare diseases and deals for instance with:

- ▶ Rett syndrome
- ▶ Huntington's disease
- ▶ Mitochondrial diseases
- ▶ Deafness
- ▶ Scoliosis
- ▶ Prader-Willi syndrome
- ▶ Angelman syndrome
- ▶ Chromosomal rearrangement
- ▶ Spastic paraplegia
- ▶ Intellectual disability
- ▶ Epilepsy
- ▶ Myopathies

The department conducts active research within national and international collaborations and utilises domestic and European Union funding. It has created the Hungarian Biobank Network. Moreover, it hosts the national patient advocacy group alliance with 33 partners. Researchers from this department participate in several European projects, such as EUROSCA or EurenOmics.

#### **g. Department of Medical Genetics, Medical School, University of Athens, Athens, Greece**

The University of Athens' Medical Genetics Laboratory constitutes a point of reference for genetics for Greece. It is the largest public centre for prenatal diagnostics in Greece offering clinical genetics, patients' evaluation, genetic counselling, and molecular genetics. It serves as a reference centre for rare diseases in Greece. The laboratory is well equipped and provides:

- ▶ Molecular genetics testing
- ▶ Molecular cytogenetics/FISH
- ▶ QFPCR, MLPA
- ▶ ArrayCGH (500 cases performed: 200 diagnosed, 300 undiagnosed)

The Department is composed of 5 faculty members, 4 post doctorate students, and 4 laboratory technicians. With its extensive experience it has made a substantial contribution to the diagnosis, management and prevention of many genetic disorders, and is recognized as a National- and European reference centre for several genetic diseases with respect to applied research and diagnostics within Southern Europe.

#### **h. Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia**

This institute was founded in 1986 and its projects are mostly focused on biotechnology processes and research related to the analysis of the genome organisation and the regulation of gene expression in different organisms.

There are 6 research units (“Laboratories”):

- ▶ Molecular biomedecine
- ▶ Molecular biology
- ▶ Human molecular genetics
- ▶ Microbial molecular genetics and ecology
- ▶ Molecular genetics of industrial microorganisms
- ▶ Laboratory for plant molecular biology

The Institute is composed of 85 staff members with 77 researchers and 8 administration and technical support members.

The Institute is associated to the International Centre for Genetic Engineering and Biotechnology (ICGEB) in Trieste, Italy (as an affiliated site) and is involved in all activities of this transnational consortium. It also collaborates for instance with Orphanet, EuroGentest, TREAT-NMD Neuromuscular Network, and the Cystic Fibrosis Mutation Database. The Institute aims to become the national centre for rare diseases in Serbia.

Examined disorders include for instance:

- ▶ Cardiomyopathies
- ▶ Thalassemia
- ▶ Haemophilia
- ▶ Neuromuscular diseases
- ▶ Gaucher’s disease
- ▶ Gorlin syndrome
- ▶ PraderWilli syndrome
- ▶ Leukemia
- ▶ Metabolic disorders

#### **i. The National Genetics Centre, Muscat, Sultanate of Oman**

The Sultanate of Oman is located along the Southeastern coast of the Arabian Peninsula. The population of Oman is currently estimated at approximately 3.4 million inhabitants. As with many communities in the Middle East, consanguineous unions are relatively common. Consequently, genetic drift has led to a concentration of genes for specific disorders in certain tribes and population groups. In the last two decades infant mortality has decreased, mainly thanks to the “Clean Water Act”. However, now it is difficult to further decrease the remaining infant mortality, which is mainly due to genetic diseases.

The Muscat Centre for instance examines:

- ▶ Thalassemia
- ▶ Down syndrome (also commonly present in younger females for unknown reasons)
- ▶ Dwarfism

- ▶ Skeletal dysplasia
- ▶ Robinow syndrome
- ▶ Spinal muscular atrophy
- ▶ Cystic fibrosis
- ▶ Osteopetrosis
- ▶ Osteogenesis imperfecta

Medical care is provided free of charge for the Omani population, with 10% of Omani families having a child with a congenital disorder, and approximately half of these also associated with intellectual disability. Infertility is rather high, which represents a problem for families and the country. The National Genetic Centre in Oman has a newly established laboratory which can perform FISH, arrayCGH and sequencing in rare diseases.

**j. Clinical Institute of Laboratory Diagnosis, Zagreb University School of Medicine and Clinical Hospital Centre, Zagreb, Croatia**

This institute was established in 1991 and deals with prenatal and postnatal diagnostics of monogenic (rare) and multifactorial diseases such as:

- ▶ Muscular dystrophy
- ▶ Spinal muscular dystrophy
- ▶ Cystic fibrosis
- ▶ Hemochromatosis
- ▶ Wilson's disease
- ▶ Gilbert's syndrome
- ▶ Hyperlipidemia
- ▶ Primary immunodeficiency

The Institute analyses approximately 2,000 samples of rare diseases per year and has a clinical unit for pharmacogenics, as well as for leukemias. Domestic and international collaboration is important part of its research activities.

**k. Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poland**

Actually, two main diseases are being studied at the Institute of Human Genetics of the Polish Academy of Sciences: cystic fibrosis and primary ciliary dyskinesia. Primary ciliary dyskinesia is generally underdiagnosed in Poland and is a highly genetically heterogeneous disorder, to which currently 26 genes have been associated. Mutations in these genes could explain 45 to 60% of all cases. The Institute has broad domestic and international collaborations. It also studies acute lymphoblastic leukemias (404 paediatric patients). Clinical syndromology and rare disease testing is carried out at other sections of this Institute.

## **I. National Coordination Centre for Rare Diseases, Charles University Prague, University Hospital Motol, Prague, Czech Republic**

In Czech Republic, there is one clinical geneticist per 100,000 individuals. Genetic services are fully reimbursed, including cytogenetics and molecular genetics. There is a national database which curates 450 DNA diagnostic tests for rare diseases in the entire country. This department is affiliated to many European research initiatives, including for instance Orphanet, Eurogentest or RD-Connect. Patients are recruited through the Clinical Genetics Department, which has long-term expertise in clinical syndromology. Czech Republic has a National plan for rare diseases, but budget for its implementation is lacking due to government austerity. There are 19 clinical geneticists and the Motol Centre belongs to one of the largest in the country performing close to 4,500 genetic tests per year for all kinds of rare diseases. DNA sequencing (Sanger/NGS), karyotyping, arrayCGH etc., are available. This centre cooperates internationally and nationally with rare disease patients' associations. Examined patients are mostly Czech, Slovak, but also increasingly come from Eastern Europe (Russia, Ukraine, Romania or Vietnam).

### **6. Mechanisms to enable international data sharing**

#### **a. Cafe Variome**

There is a considerable need and desire for networks of diagnostic laboratories/disease consortia to be able to check each others' databases for the presence of mutations they observe in their own patients. But this is countered by the understandable reluctance and impracticality of sharing the content of each group's database with other labs, or indeed with the wider world. Cafe Variome (<http://www.cafevariome.org/>) has been developed in order to enable the 'open discovery' of data (rather than data sharing) between networks of diagnostics laboratories or disease consortia that know/trust each other and share an interest in certain causative genes or diseases. The existence of data is shared, rather than the substance (i.e. exact variants). It enables laboratories to share the fact that a variant has been seen in a patient without necessarily revealing the underlying data, thus overcoming legal issues related to patient confidentiality. Cafe Variome is reaching out to groups and projects that could benefit from such a system, allowing them and their fellow laboratories to securely share data amongst themselves.

At the moment, the Cafe Variome network collaborators are:

- ▶ DMuDB
- ▶ Denmark diagnostic network
- ▶ German diagnostic network
- ▶ French diagnostic network (Interactive Biosoftware)
- ▶ Belgium diagnostic network
- ▶ Canadian diagnostic network (FORGE Canada/Care4Rare)
- ▶ Netherlands diagnostic network
- ▶ Ehlers Danlos Syndrome network
- ▶ Collagen disease network
- ▶ Inherited Colon Cancer network (InSiGHT)

Once datasets have been discovered, data could be subsequently accessed under three different models. Cafe Variome is not a database. It is a research menu for what exists in various data sources. It is designed to enable users to ask the question "where can certain data be found?", seeking to access those data under one of the three models listed below, as stipulated by the data owner/submitter:

- ▶ Open Access: it is like a journal, where variant records are publicly available for access
- ▶ Restricted Access: the user requests access to variant from the data owner/submitter, the user must either belong to a pre-approved group or must request access from the data owner in order to access the variant record(s)
- ▶ Linked Access: it only reports the existence of a variant and the user is linked to the source database in order to access the full record

Access control in the Cafe Variome interface allows data owners to modify permissions for whole datasets down to the individual variant level. It requires only three input fields per recorded mutation: HGVS name, gene name & reference URL.

Cafe Variome is affiliated with the University of Leicester (UK), GEN2PHEN.org, and PhenoSystems.com and has received funding from the European Community's Seventh Framework Programme.

## **b. PhenoTips**

Phenotypic descriptions are evocative for humans, but not for computers. Most genotype/phenotype databases store little about phenotypes. Ontologies are required as they represent standardized and defined terms with recognized synonyms. The Human Phenotype Ontology (HPO) contains over 10,000 terms to describe human abnormalities.

PhenoTips (<http://phenotips.org/>) is a free open source software tool for collecting and analyzing phenotypic information based on HPO for patients with genetic disorders. It can be used by geneticists and paediatricians (if they use genetics in their practice). The user interface closely mirrors clinician workflows so as to facilitate the recording of observations made during the patient encounter. This easy-to-use front-end, compatible with any device that runs a Web browser, is coupled with a standardized database back-end where phenotypic information is represented using the HPO. It supports abbreviations and orthographic errors.

Collected data include demographics, medical history, family history, physical and laboratory measurements, physical findings, and free-form notes.

In addition to data collection, PhenoTips automatically analyzes a wide range of measurements and plots live the corresponding growth curves. It also supports clinical diagnosis based on the entered data, providing possible diagnoses based on searches of OMIM with the HPO terms. It was suggested during the Workshop that PhenoTips could be directly linked to ORPHA codes of Orphanet. In the future, a search for genes will be added to PhenoTips.

PhenoTips is developed in the Computational Biology Lab at University of Toronto's Computer Science Department, in collaboration with geneticists and clinicians from SickKids and CARE4RARE.

## **c. PhenomeCentral**

The availability of low-cost genome sequencing has allowed for the identification of the molecular cause of hundreds of rare genetic disorders. But because the discovery of disease-causing variants requires confirmation of the mutation or gene in multiple unrelated individuals, an even larger number of genetic

disorders remain unsolved due to difficulty identifying second families. It is therefore critical to establish effective and secure data-sharing techniques that allow clinicians and scientists to identify additional families via phenotype and genotype searches.

To address this need, PhenomeCentral (<http://phenomecentral.org>) has been developed. It is a repository for secure data sharing targeted to the rare disorder community. Each patient record within PhenomeCentral consists of a phenotypic description capturing observed abnormalities as well as relevant absent manifestations, expressed HPO terms. Furthermore, each record can be labelled by the creator as:

- ▶ Private: hidden from everyone except the contributor
- ▶ Public: viewable and searchable by all registered users
- ▶ Matchable: the record cannot be directly viewed or searched, but is reachable via an automated phenotype matching system (following Cafe Variome principles) which informs contributors of the existence of profiles similar to their cases. The phenotypic features shared among these records are presented without revealing additional patient information or the contributors, enabling direct communication for any subsequent data sharing

PhenomeCentral currently incorporates phenotype data for more than 400 patients with rare genetic disorders without a molecular diagnosis, including: 200 from the Canadian CARE4RARE project and 150 from the NIH Undiagnosed Diseases Program. Clinical geneticists and scientists studying rare disorders can request accounts, and new patients can be added either using the PhenoTips User Interface, built into PhenomeCentral, or uploaded in bulk.

## 7. Next steps

- ▶ Link PhenoTips to ORPHA Codes of Orphanet
- ▶ Develop collaborations between the different institutes present at the workshop
- ▶ CARE4RARE ([www.care4rare.ca](http://www.care4rare.ca)) is open to international collaborators and interested clinicians can submit interesting patients/families for study (exome sequencing) using the link on the website. Control of the project remains with the submitter and CARE4RARE will facilitate data analysis as needed. Interested groups should contact [thartley@cheo.on.ca](mailto:thartley@cheo.on.ca)

