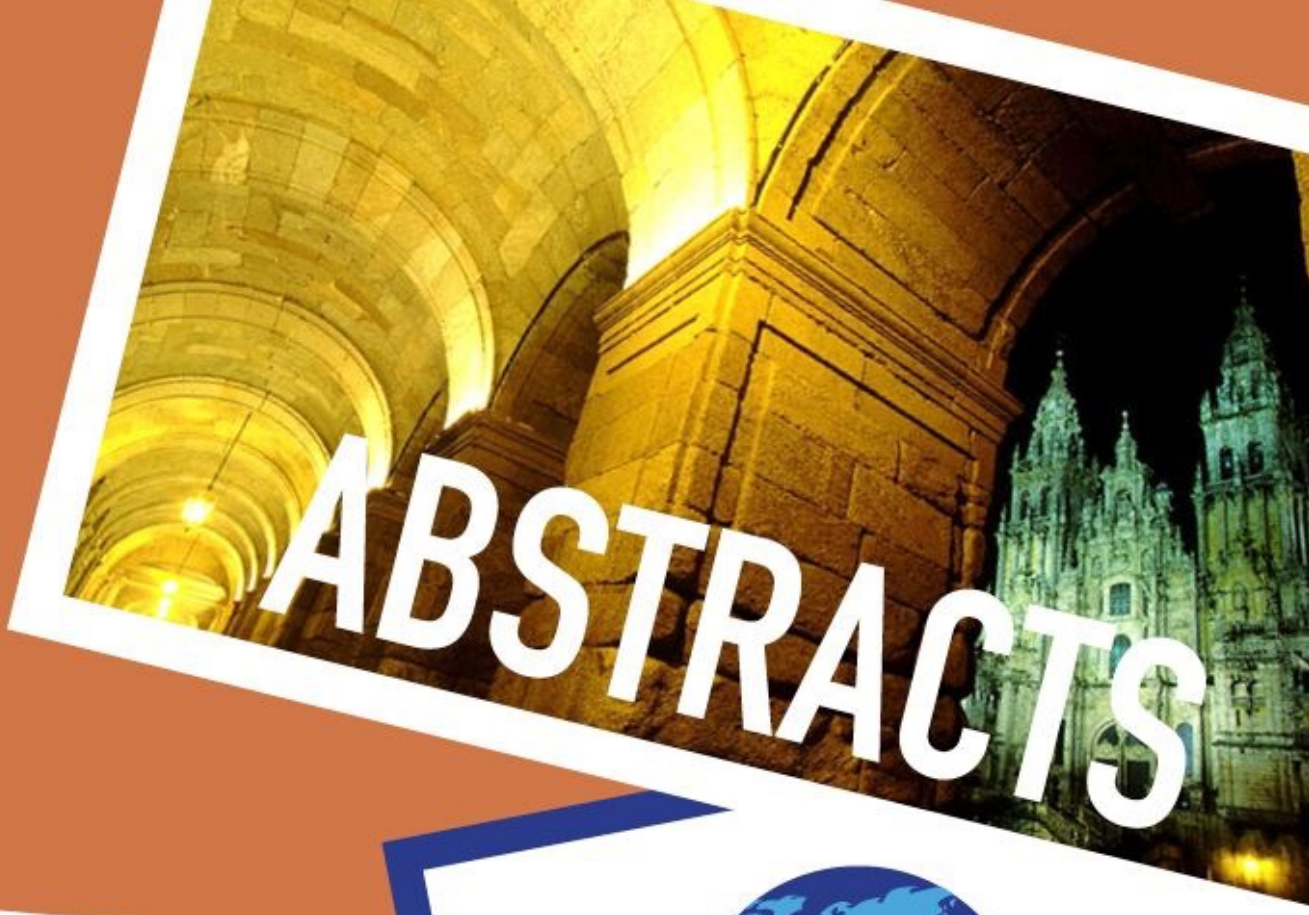

ABSTRACTS

THE HUMAN VARIOME PROJECT

an NGO official partner of UNESCO

sharing data · reducing disease

5 - 7 June 2017
Santiago de Compostela
Spain

14th International Symposium on Variants in the Genome

THE HUMAN VARIOME PROJECT
an NGO official partner of UNESCO
sharing data · reducing disease

14th International Symposium on Variants in the Genome: detection, sequencing & interpretation

5 - 7 June 2017

NH Collection Santiago
Santiago de Compostela, Spain

Scientific Programme Committee

Prof. Johan T. den Dunnen (Leiden, Nederland) CHAIR
Prof. Sir John Burn (Newcastle, UK)
Prof. Angel Carracedo (Santiago de Compostela, Spain)
Dr Reece Hart (San Francisco, CA, USA)
Dr Andreas Laner (Munich, Germany)
Dr Maria-Jesus Sobrido (Santiago de Compostela, Spain)

Organising Committee

Maria Torres (CEGEN-PRB2, Universidade de Santiago de Compostela)
Rania Horaitis; Event Manager (Meeting Makers www.meeting-makers.com)

Previous Meetings

1991 Oxford, UK
1993 Lago D'Orta, Italy
1995 Visby, Sweden
1997 Brno, Czech Republic
1999 Vicoforte, Italy
2001 Bled, Slovenia
2003 Palm Cove, Australia
2005 Santorini, Greece
2007 Xiamen, China
2009 Paphos, Cyprus
2011 Santorini, Greece
2013 Lake Louise, Canada
2015 Leiden, Netherlands

Monday 5th June

8.30 - 14.00 Exhibitor Bump In

8.00 - 11.00 **REGISTRATION**
Hotel Foyer

9.00 - 10.40 **PRE-MEETING WORKSHOP 1**
VARIANT NOMENCLATURE
Room: Allariz + Noia
Johan T. den Dunnen
Leiden Univ. Medical Center, Leiden, Netherlands

PRE MEETING WORKSHOP 2
CLINICAL GENOMICS – INTERPRETATION & REPORTING WITH QIAGEN BIOINFORMATICS
Room: Obradoiro
Ruth Burton
Advanced Genomics, Qiagen

10.40 - 11.00 Coffee Break
Room: Obradoiro Terrace

11.00 - 11.10 **INTRODUCTION & WELCOME**

Introduction
Johan T. den Dunnen
Leiden Univ. Medical Center, Leiden, Netherlands

Welcome
Maria-Jesus Sobrido & Angel Carracedo
Instituto de Investigaciones Sanitarias de Santiago / Fundación Pública Galega de Medicina Xenómica

11.10 - 13.00 **PLENARY SESSION 1 – VARIANT CALLING AND ANNOTATION FOR DIAGNOSTIC APPLICATIONS**
Room: Obradoiro
Chair: Johan T. den Dunnen

11.10 - 11.50 **Phenotype Driven Genomic Diagnostics**
Peter Robinson
Computational Biology Group at the Jackson Lab. for Genomic Medicine, Farmington CN, USA

Monday 5th June

11.50 - 12.20	**CNV detection from targeted next-generation panel sequencing data in routine diagnostics** Anna Benet-Pagès *Medizinisch Genetisches Zentrum, Munich, Germany*
	PRESENTATIONS FROM ABSTRACTS
12.20 - 12.40	**GeneHancer and VarElect: disease interpretation of whole genome sequence variants** Doron Lancet *Weizmann Institute of Science, Rehovot, Israel*
12.40 - 13.00	**ChromiumTM: Full spectrum genome analysis with Linked-Reads** Steve Giavas *10x Genomics Inc.* *10x Genomics is a sponsor of the 14th International Symposium on Variants in the Genome*
13.00 - 14.00	Lunch Room: Azabache Restaurant
14.00 - 15.30	**PLENARY SESSION 2 – THE BRCA CHALLENGE** **Room: Obradoiro** Chair: Stephen Chanock
14.00 - 15.00	**The BRCA Challenge** Prof. Sir John Burn *Institute of Genetic Medicine, International Centre for Life, Newcastle upon Tyne, UK* Gunnar Ratsch *ETH Zurich, Memorial Sloan-Kettering Cancer Center, Weill Cornell Medical College*

Monday 5th June

15.00 - 15.30	**BRCA Exchange Mobile: Enabling Patient Access, Notification, and Case-Level Data Ingress** Faisal Alquaddoomi *ETH Zurich, Switzerland*
15.30 - 16.10	**Poster Session 1** Room: Quintana & Lobby Odd numbers to present: POS001, POS003 etc
16.10 - 16.20	Assemble for Group Photograph **Room: Obradoiro Terrace**
16.20 - 18.00	PLENARY SESSION 3 – The BRCA Challenge (continued) **Room: Obradoiro** Chair: Prof. Sir John Burn
16.20 - 16.50	**The Genomics England 100,000 Genomes Project: Establishing a centralised national resource of genomic data** Claire Turnbull *Genomics England, Queen Mary University of London and Institute of Cancer Research, London, UK*
16.50 - 17.20	**Novel genes involved in Fanconi anemia, DNA repair and cancer predisposition: the clinical relevance of functional studies of genetic variants** Jordi Surrallés Calonge *Universitat Autònoma de Barcelona, Catalunya, Spain*

Monday 5th June

PRESENTATIONS FROM ABSTRACTS

17.20 - 17.40	**A unified framework for prioritization of variants of uncertain significance In hereditary breast and ovarian cancer (HBOC)** Peter Rogan *University of Western Ontario and CytoGnomix Inc., London, Canada*
17.40 - 18.00	**Pilot multi-gene testing in Hereditary Breast-Ovarian Cancer** Paola Carrera *Irccs San Raffaele Scientific Institute, Milano, Italy*
19.00 - 20.30	**WELCOME RECEPTION** **Room: Obradoiro Terrace**

Tuesday 6th June

8.30 - 10.30	**PLENARY SESSION 4 – VARIANT INTERPRETATION IN THE CLINIC** **Room: Obradoiro** Chair: Reece Hart
8.30 – 9.00	**A systematic framework for the clinical interpretation of chromosomal copy number variants** Swaroop Aradhya *On behalf of the ACMG/ClinGen Structural Variant Working Group*
9.00 - 9.30	**Challenges in Variant Interpretation – How to minimize inter and intra-laboratory inconsistencies** Andreas Laner *Medizinisch Genetisches Zentrum, Munich, Germany*
9.30 - 10.00	**From a list of variants to a diagnostic report: extracting clinically relevant information** María-Jesús Sobrido *Instituto de Investigaciones Sanitarias de Santiago / Fundación Pública Galega de Medicina Xenómica*
10.00 - 10.30	**Accessing the full size-spectrum of human genetic variation using PacBio long-read SMRT sequencing on the Sequel System** Luke Hickey *Senior Director of Human Biomedical Sciences, PacBio* *PacBio is a Sponsor of the 14th International Symposium on Variants in the Genome*
10.30 - 11.00	Coffee Break Room: Quintana & Lobby

Tuesday 6th June

11.00 - 13.10 **PLENARY SESSION 5 – POPULATION GENETICS & FORENSIC APPLICATIONS**

Room: Obradoiro

Chair: Maria-Jesus Sobrido

11.00- 11.30 **Challenges with the compilation and naming of new variation revealed by massively parallel sequencing of forensic markers**

Christopher Phillips

University of Santiago de Compostela, Galicia, Spain

11.30 - 12.00 **Naming Genetic Variation in Forensic Science: Alignment and Nomenclature of Next Generation Sequence Alleles**

Walther Parson

President International Society of Forensic Genetics, Institute of Legal Medicine, Innsbruck, Austria

PRESENTATIONS FROM ABSTRACTS

12.00 - 12.20 **BaseSpace Variant Interpreter: A new platform to improve the speed of genomic interpretation and facilitate collaborative knowledge sharing**

Jennifer Harrow

Programme Manager Population Sequencing, Illumina Inc., UK

12.20 - 12.40 **Safe variant annotation sharing across laboratories**

Beat Wolf

HES-SO, Informatics, Fribourg, Switzerland

12.40 - 13.10 **LGC SPONSORED PRESENTATION**

Variant detection and the challenges beyond: what is needed to implement genetic information for clinical use?

Prof. Dr. Daniela Steinberger

Human Geneticist, Medical Director, bio.logis Center for Humangenetics, Frankfurt am Main, Germany

Tuesday 6th June

13.10 - 14.10	Lunch Break Room: Azabache Restaurant
14.10 - 15.40	**SPONSORED WORKSHOP – ALAMUT** **Room: Obradoiro** André Blavier & Séverine Lair *Interactive Biosoftware*
15.40 - 16.20	**Poster Session 2** Room: Quintana & Lobby Even numbers to present: POS002, POS004 etc.
16.20 - 18.30	**WORKSHOP – GENETIC VARIANT INTERPRETATION** **Room: Obradoiro** **ORGANIZERS:** • Andreas Laner • Reece Hart • Maria—Jesus Sobrido
18.30	**DAY END – EVENING AT LEISURE**

Wednesday 7th June

8.30 - 10.30	**PLENARY SESSION 6 – BIOINFORMATICS AND BIG DATA** **Room: Obradoiro** Chair: Angel Carracedo
8.30 - 9.00	**Challenges in bioinformatics for genetic diagnosis** Joaquin Dopazo *Fundacion Progreso y Salud, Clinical Bioinformatics Research Area, Sevilla, Spain*
9.00 - 9.30	**A community-developed data model for representing sequence variation** Reece Hart *Invitae, San Francisco, USA*
9.30 - 10.00	**KEYNOTE SPEAKER** **Meiosis, recombination, and the origin of a species** Peter Donnelly *Wellcome Trust Centre for Human Genetics, Oxford, UK*
10.00 - 10.30	**Raising the bar in NGS diagnostics: Challenges in Variants Identification** Zelie Dubreucq *Subject Matter Expert, at Sophia Genetics* *Sophia Genetics is a Sponsor of the 14th International Symposium on Variants in the Genome*
10.30 - 11.00	Coffee Break Room: Quintana & Lobby
11.00 - 14.00	**PLENARY SESSION 7 – NEW GENOMIC TECHNOLOGIES** **Room: Obradoiro** Chair: Andreas Laner

Wednesday 7th June

11.00 - 11.30	**Strategies for assembling high quality genome sequences** Ivo Gut *Centro Nacional de Análisis Genómico, CNAG-CRG, Barcelona, Spain*
11.30 - 12.00	**Latest applications of innovative technologies in our research and diagnostics** Johan T. den Dunnen *Leiden Univ. Medical Center, Leiden, Netherland*
	PRESENTATIONS FROM ABSTRACTS
12.00 - 12.20	**Accelerating diagnosis of hereditary diseases analytically with a knowledge network and the ACMG guidelines** Ruth Burton *Qiagen, Redwood City, USA*
12.20 - 12.50	**Rapid Genome Wide Mapping at the Single Molecule Level Using Nanochannel Arrays for Structural Variation Analysis and de novo Assembly** Sven Bocklandt *Senior Application Specialist, BioNano Genomics* *BioNano Genomics is a Sponsor of the 14th International Symposium on Variants in the Genome*
12.50 - 13.50	Lunch Break Room: Azabache Restaurant
13.50 - 15.20	**PLENARY SESSION 8 – PERSONALISED MEDICINE & PHARMACOGENOMICS** **Room: Obradoiro** Chair: Peter Donnelly
13.50 - 14.20	**Advancing Drug Discovery – Translating Small & Big Data Into Insight** Guna Rajagopal *Global Head of Computational Sciences within Discovery Sciences* *Janssen Pharmaceuticals R & D, Spring House, PA, USA*

Wednesday 7th June

14.20 - 14.50	**Innopharma: the Pharmacogenomics platform of the University of Santiago de Compostela** Mabel Loza/ Ángel Carracedo *University of Santiago de Compostela, Galicia, Spain*
14.50 - 15.20	**Advances in Personalised Medicine – Liquid Biopsy** Alexander Sartori *Manager Applications Development Agena BioScience* *Agena BioScience is a Sponsor of the 14th International Symposium on Variants in the Genome*
15.20 - 15.50	Coffee Break Room: Quintana & Lobby
15.50 - 17.00	**PLENARY SESSION 9 – PRESENTATIONS FROM SELECTED ABSTRACTS** **Room: Obradoiro** Chair: Ivo Gut
15.50 - 16.10	**PacBio long read sequencing for improved resolution of complex genomic variation** Henk Buermans *Leiden Genome Technology Center, Leiden, Netherlands*
16.10 - 16.30	**Tracing mitochondrial mutations in 3D in primary tumor, lymph node and liver metastasis** Per Ekstrom *The Norwegian Radium Hospital, Oslo, Norway*
16.30 - 16.50	**The TP53 mutation database: a paradigm for the analysis of cancer genes** Thierry Soussi *Karolinska Institutet, Stockholm, Sweden*
16.50- 17.00	**SUMMARY – CLOSING** *Johan T. den Dunnen, Maria Jesus Sobrido, Angel Carracedo*
17.00	**MEETING END**

Oral Presentation Abstracts

SESSION 1

Phenotype Driven Genomic Diagnostics

Peter N Robinson

Professor of Computational Biology, The Jackson Laboratory for Genomic Medicine, 10 Discovery Drive, Farmington, CT 06032, USA
peter.robinson@jax.org
www.jax.org
Robinson lab: https://robinsongroup.github.io/

The analysis of phenotype plays a key role in clinical practice and medical research but is difficult to compute“ on. The Human Phenotype Ontology (HPO) provides a standardized, controlled vocabulary that allows phenotypic information to be described in an unambiguous fashion and has been adopted by a number of groups including the NIH Undiagnosed Diseases Program (UDP) and the UK 100,000 Genomes Project. The HPO makes clinical data “computable”, and allow a shift from binary phenotype analysis (patient cohort vs. control) to an analysis of phenotypic profiles that offers a foundation for computational approaches to integrating clinical data in precision medicine. HPO-based algorithms integrate phenotype specificity, imprecision, noise and frequency to identify matching diseases and patients. We have developed Exomiser to identify promising candidate genes in whole exome sequencing WES studies by ranking candidate genes according to phenotypic similarity to human, mouse, and zebrafish mutant phenotypes. Using simulated exomes and the NIH UDP patient cohort showed Exomiser ranked the causal variant as the top hit in 97% of known disease–gene associations. We have developed a machine learning algorithm that provides a pathogenicity score for each base of the non-coding genome, and extended the Exomiser to include an assessment of regulatory regions, in an application called the Genomiser. Simulations show an ability to rank the seeded causative variant in first place over an entire genome in over 60% of cases. In this talk, I will review how computational phenotype analysis with the HPO works and how we have used it for the Exomiser and Genomiser.

CNV detection from targeted next-generation panel sequencing data in routine diagnostics

Anna Benet-Pagès, Anke M. Nissen, Janine Graf, Christina Rapp, Melanie Locher, Andreas Laner, Elke Holinski- Feder.

Medizinisch Genetisches Zentrum, MGZ, Munich, Germany

Gene dosage abnormalities account for a significant proportion of pathogenic mutations in rare genetic disease related genes. In times of next generation sequencing (NGS), a single analysis approach to detect SNVs and CNVs from the same data source would be of great benefit for routine diagnostics. However, CNV detection from exon-capture NGS data has no standard methods or quality measures so far. Current bioinformatics tools depend solely on read depth which is systematically biased. We developed a novel approach based on: 1. utilization of five independent detection tools to increase sensitivity, 2. different reference sets for different kits and normalization against samples from the same sequencing run to improve robustness against workflow conditions, 3. definition of special quality thresholds for single exon events to minimize false negatives, 4. identification of reliable regions by assessment of capture efficiency using a reference set of CNV negative patients to minimize false positives. A CNV is called in a reliable region if at least two out of five tools are concordant for the respective CNV. The

pipeline shows a sensitivity of 80% and a precision of 95%. Within routine gene panel diagnostics we analyzed a total of 1088 patients indicated to have rare Mendelian diseases for SNV and CNVs. In 32 patients a CNV was detected in genes associated with the respective individual phenotype. Interestingly, in several cases the CNV completed the patients report as it was detected in genes with a recessive mode of inheritance where previously only a heterozygous pathogenic SNV was found. Overall, with the additional analysis of CNVs we increased the diagnostic yield from 15% (class 4, 5 single nucleotide events) to 18%. However, there are still issues in the detection of CNVs from NGS data for routine diagnostics. CNV pipelines are very prone to errors caused by enrichment inconsistencies compared to SNV detection tools. The assessment of sensitivity and specificity is difficult due to the lack of datasets to validate CNV detection pipelines. Originally, the analysis of CNVs was performed mainly in patients with mental retardation disorders, resulting in a paucity of CNV data linked to other Mendelian diseases. Moreover, the identification of the actual size and thus the assessment of pathogenicity of a CNV is difficult, because targeted NGS gene panels do not cover all genes. In conclusion, NGS data is a suitable data source for the simultaneous detection of SNVs and CNVs for clinical diagnosis; however, with the current tools it is only applicable in accurately validated regions.

GeneHancer and VarElect: disease interpretation of whole genome sequence variants

Simon Fishilevich, Naomi Rosen, Michal Twik, Rotem Hadar, Tsippi Iny-Stein, Marilyn Safran and Doron Lancet

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7610001, Israel

Purpose: The emergence of whole genome sequencing (WGS) poses considerable challenges to variant disease interpretation. A typical WGS of an individual identifies ~5million non-reference variants, a 50 fold increase compared to whole exome sequencing (WES). A considerable proportion of this “variant avalanche” (10-15%) resides within transcription regulatory elements - promoters and enhancers. Promoters are relatively easy to identify due to their stereotyped positioning at the immediate 5’ neighborhood of genes, and their targets are quite obvious. In contrast, the identification of enhancers constitutes a major challenge. An equally difficult task is identifying the connections between these distant-acting regulatory elements and their target genes. Genomic enhancers are centrally involved in the spatiotemporal orchestration of gene expression in embryonic development and in cell differentiation. This makes them prime novel targets for annotating the plethora of non-coding variants in WGS, and interpreting them in the realms of health and diseases.

Methodology: We created GeneHancer, a novel regulatory element database, in the framework of the GeneCards suite (www.genecards.org). We integrated four enhancer data sources: a) 176,000 enhancer regions from the ENCODE project (https://www.encodeproject.org/); b) 213,000 elements from the Ensembl regulatory build (http://www.ensembl.org/index.html); c) 43,000 elements from FANTOM (http://fantom.gsc.riken.jp/), identified via enhancer RNAs (eRNAs); d) 1,700 experimentally-validated elements from the VISTA enhancer browser (http://enhancer.lbl.gov/). Subsequently, we consolidated gene-enhancer links obtained by five methodologies: a) GTEx expression quantitative trait loci (eQTLs, http://www.gtexportal.org/home/), b) Capture Hi-C promoter-enhancer long range interactions (PMID 25938943); c) FANTOM expression correlations between eRNAs and candidate target genes; d) Expression correlations between enhancer-targeted transcription factors and genes; e) Enhancer-gene genomic distance scores.

Results: GeneHancer portrays 285,000 integrated non-redundant candidate enhancers (covering 12.4% of the genome), along with annotation-derived confidence scores. In parallel, our database incorporates ~1.02 million integrated and scored gene-enhancer links involving 101,337 genes. Among these, we define a subset of "double elite" enhancer-gene pairs, based on the conjunction of two or more methods for both entities. This allows WGS variants within enhancers to be interpreted with high confidence, based on high-probability target gene links. These WGS analysis capabilities are being embedded within the GeneCards Suite, among others by modifying VarElect and TGex, its next generation sequencing (NGS) disease interpretation tools [PMID: 27357693].
For WES, VarElect prioritizes a list of variant-containing genes by seeking the relevance of such genes to phenotype/disease/symptom keywords, as inferred from the comprehensive web-mined information within the GeneCards knowledgebase. For WGS, TGex maps non-coding variants onto enhancers, using GeneHancer tables. Then, a modified VarElect assigns GeneHancer target genes to such enhancers. These genes are added to VarElect's input gene list for performing disease interpretation. Enhancer variants are ranked by a combination of phenotype scores and GeneHancer scores.

Conclusion: The combination of GeneHancer and VarElect, along with the power of the GeneCards suite's comprehensive gene and disease information, provides a facile, automatable route to discovering the genic roots of diseases, including in clinical projects with extremely large genome counts.

Support: LifeMap Sciences grant

ChromiumTM: Full spectrum genome analysis with Linked-Reads

Steve Glavas, Sarah Garcia, Claudia Catalanotti, Haynes Heaton, Patrick Marks, Michael Schnall-Levin, Stephen Williams, Andrew Wei Xu, Grace Zheng, Deanna M. Church

10x genomics, Inc.

High-throughput sequencing has revolutionized genome analysis. However, it is clear that traditional short read methods provide an incomplete view of the genome and result in an incomplete understanding of the clinical and biological complexity present. In particular, the lack of long-range information combined with inherent limitations in the mapping of short reads severely limits the robust identification of structural variants, haplotypes, and variants in difficult regions of the genome. To address these problems, we developed the Chromium platform, a technology that retains long-range information while maintaining the power, accuracy, and scalability of short read sequencing. At its core, haplotype-level dilution of long input molecules into >1 million barcoded partitions creates a novel data type referred to as 'Linked-Reads' that enables high-resolution genome analysis with minimal DNA input (~1 ng).

Coupling Linked-Reads with novel algorithms that take advantage of these linkages allows for improved haplotype reconstruction without sacrificing variant calling accuracy. We obtain 99% SNP sensitivity with greater than 99% PPV in confident regions for genome and exome on the NA12878 sample. We see phase blocks with an N50 of 4.1 Mb for genomes and 222 Kb for exome. We also show improved performance in regions of the genome typically inaccessible due to the presence of paralogous sequence/highly repetitive regions. Highly homologous sequences traditionally leading to ambiguous mapping can now be analyzed when

associated with distinct barcodes. We estimate that we can rescue ~50 Mb of previously inaccessible sequence. We demonstrate that clinically-relevant variants can be identified in these previously inaccessible regions in genes including *CYP2D6*, *SMN1*, and *STRC*.

Finally, retention of long range information facilitates the identification of copy number variants, copy neutral inversions, inter- and intra-chromosomal events, and more complex structural rearrangements. In a set of 22 cell lines with known clinically relevant structural events including copy number altering and neutral events, linked-read genome sequencing was able to identify 20/22 known variants. Both of the events that were not identified were balanced translocations- one was called as a low quality candidate variant, and the other involves the pericentromeric region of chr16, a known gap region in GRCh37. Even at reduced sequencing coverage (5-50Gb), these events are clearly identifiable. The utilization of Linked-Reads significantly improves our ability to reconstruct individual haplotypes and provides confidence in variant calls by providing an additional data type- phasing.

SESSION 2

The BRCA Challenge

Gunnar Ratsch, Lena Dolman, Stephen Chanock, and Sir John Burn on behalf of the BRCA Challenge Steering Committee, Global Alliance for Genomics and Health (GA4GH) and the Human Variome Project

Diagnosis of disease predisposition due to pathogenic genetic variants is hampered by the lack of a unified curated resource. Over the last decade, efforts to develop curated databases, supported by the Human Variome project, have seen some success. The more recent emergence of large scale genomic data sets prompted the launch of the Global Alliance for Genomics and Health (GA4GH) to tackle the core challenges of data integration. As a demonstration project, the BRCA Challenge was launched in 2014, based on the high public profile and frequent diagnostic testing of these two genes. To advance understanding of the genetic basis of hereditary breast, ovarian, and other cancers, and to better enable accurate clinical care, this project proposed to: i) share *BRCA1/2* variants publically; ii) create a curated list of *BRCA1/2* variants, interpreted by expert consensus from the ENIGMA Consortium, and; iii) create an environment for collaborative sharing of phenotypic data. In 2016, a web portal to support these goals, the "BRCA Exchange" (brcaexchange.com), was launched. Today, the BRCA Exchange includes over 18,000 unique, de-duplicated variants, making it the largest public source for *BRCA1/2* variant information. Using the GA4GH Genomics API, this portal aggregates variant data from public sources (ClinVar, LOVD, Ex-UV Multifactorial Database, 1000 Genomes, ExAC, BIC, ENIGMA, and ESP) with automated monthly updates. The portal consists of two tiers: the 'Expert-Reviewed' tier, which supports clinical care by displaying consensus expert classifications (and associated evidence) for almost 5,000 variants as assigned by the ENIGMA Consortium; and the 'All Public Data' research tier, which displays a range of evidence (including allele frequencies, predictive algorithms, and assertions of pathogenicity) for the entire dataset, derived from the original submitters. The full genomic dataset is downloadable and versioned, enabling research and improved speed and efficiency for curators. The BRCA Exchange ultimately enables cyclical knowledge exchange among researchers, care providers, and diagnostic labs, allowing for improved patient management in the hereditary cancer domain.

BRCA Exchange Mobile: Enabling Patient Access, Notification, and Case-Level Data Ingress

Faisal Alquaddoomi (ETH Zurich, Switzerland), Prof. Gunnar Rätsch

ETH Zurich, Switzerland

The BRCA Challenge was established in March 2014 at the first plenary meeting of the GA4GH, and soon after the BRCA Exchange, a web portal for distributing information about BRCA variants to the public, was begun in June 2015. The subject of this talk, the BRCA Exchange mobile app (tentatively titled "BRCA Exchange Mobile"), builds upon the BRCA Exchange infrastructure by providing a "mobile-first" interface to the database. It also provides capabilities for keeping patients in the loop as variants are updated, and will eventually allow ingress of patient data into the BRCA Exchange ecosystem.

The app's development is divided into two phases, the first of which is nearing completion and the second which is underway. In phase I, we focus on database-to-user flow — specifically the patient can access the database via the app and "follow" variants of interest. When the database is updated, the patient receives push notifications on their mobile device. In phase II, we will focus on patient-to-database flow. This phase adds the ability for patients to scan and upload paper reports or complete on-device questionnaires, which are then uploaded to a secure location for anonymization and aggregation. The anonymized and aggregated case-level data will be disbursed to researchers and, after analysis, be integrated into the public database, thus closing the loop.

Acknowlegements: Baroness Morgan of Drefelin (Breast Cancer Now, Cancer Research UK); Baylor College of Medicine (Houston, USA); Curie Institute (Paris, France); Dana-Farber Cancer Institute (Boston, USA); Duke University (Durham, USA); ETH Zurich (Zurich, Switzerland); Harvard Medical School (Boston, USA); Hunstman Cancer Institute (Salt Lake City, USA); Institute of Cancer Research (London, UK); Leiden University Medical Center (Leiden, Netherlands); Mayo Clinic (Rochester, USA); McGill University (Montreal, Canada); Memorial Sloan Kettering Cancer Center (New York, USA); National Cancer Institute (Rockville, USA); National Human Genome Research Institute (Bethesda, USA); National Institute of Health (Bethesda, USA); Newcastle University (Newcastle upon Tyne, UK); Osaka University (Osaka, Japan); QIMR Berghofer Medical Research Institute (Herston, Australia); UC Santa Cruz Genomics Institute (Santa Cruz, USA); University of California San Francisco (San Francisco, USA); University of Cambridge (Cambridge, UK); University of Pennsylvania (Philadelphia, USA)

SESSION 3

The Genomics England 100,000 Genomes Project: Establishing a centralised national resource of genomic data

Dr Clare Turnbull, MD PhD MA MSc MRCP MFPH
Clinical Reader in Genomic Medicine

Genomics England, Queen Mary University of London, nstitute of Cancer Research, London, Guys and St Thomas NHS Foundation Trust, London

The Genomics England 100,000 Genomes Project was launched in order to develop the infrastructure and introduce the processes and pathways necessary for implementation of whole genome sequencing across the National Health Services of the UK. Focusing primarily on rare disease and cancer, patient recruitment began in 2014. We have now established a national sequencing centre as

well as a national data architecture for collection of clinical data, storage and processing genomic data and presentation back of analysed genomes alongside the clinical and longitudinal life-course data to the clinical and research communities. Alongside the program, we have launched a national program of education in genomics targeted at healthcare employees, including Masters Courses at 10 universities. To facilitate engagement and acceptability to the public, we have a comprehensive patient participant involvement around the program, as well as developing resources for public education about genomics. I shall update on the progress made and challenges experienced within the UK 100,000 Genomes Project, focusing on our systems for storage of variant data and approaches to data federation.

Novel genes involved in Fanconi anemia, DNA repair and cancer predisposition: the clinical relevance of functional studies of genetic variants

Jordi Surrallés

Genetics Service, Hospital de Sant Pau, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, and CIBER on Rare Diseases, Barcelona, Spain.

The biomedical relevance of genome maintenance is illustrated by the severe clinical consequences of mutations in DNA repair genes. A clear example are genes involved in the repair of DNA interstrand-cross links and double strand breaks by homologous recombination (HR), such as BRCA1, BRCA2, PALB2, BRIP1, or RAD51C. Mutations in these genes cause familial breast cancer and Fanconi anaemia (FA) in monoallelic or biallelic carriers, respectively. Furthermore, the proteins encoded by many of these genes are crucial for the modulation of the response of cancer cells to chemotherapeutics, including cisplatin and PARP inhibitors. Therefore, the identification of additional components of this DNA repair pathway is of extreme biomedical importance. By using whole exome sequencing and interactomics we have recently identified and functionally studied novel components of this pathway and uncovered their association to breast and colon cancer susceptibility. Functional studies of variants identified in these novel genes are of critical importance to understand their role in cancer predisposition or FA-related bone marrow failure and to correlate the genetic variant with the clinical outcome. The correct genetic characterization of patients with FA is essential for developing therapies, including hematopoietic stem cell transplantation from a savior sibling donor after embryo selection, untargeted gene therapy, targeted gene therapy, or genome editing using genetic recombination or engineered nucleases. Newly acquired knowledge about FA promises to provide a cure in the near future.

A Unified Framework For Prioritization Of Variants Of Uncertain Significance In Hereditary Breast And Ovarian Cancer (HBOC)

Eliseos Mucaki[1], Natasha Caminsky[1], Ruipeng Lu[1], Joan Knoll[1,2] and Peter Rogan[1,2].

[1]University of Western Ontario, [2]CytoGnomix Inc.

Purpose: A significant proportion of HBOC patients receive uninformative genetic testing results, an issue exacerbated by the overwhelming quantity of variants of uncertain significance identified. We apply information theory (IT) to predict and analyze non-coding variants of uncertain significance (VUS) in regulatory, coding, and intronic regions based on changes in binding sites in these genes. This provides a unifying framework where, aside from protein coding changes, pathogenic variants occurring within sequence elements can be prioritized based

on their recognition by proteins involved in mRNA splicing, transcription, and untranslated region binding and structure. To support the utilization of IT analysis, we established IT-based variant interpretation accuracy by performing a comprehensive review of mutations altering mRNA splicing in rare and common diseases[1].

Methods: We captured and enriched for coding and non-coding variants in genes known or suspected to increase HBOC risk. Custom oligonucleotide baits spanning the complete coding, non-coding, and intergenic regions 10 kb up- and downstream of *ATM, BRCA1, BRCA2, CDH1, CHEK2, PALB2, TP53, ATM*, *BARD1*, *BRCA1*, *BRCA2*, *CDH1*, *CHEK2*, *EPCAM*, *MLH1*, *MRE11A*, *MSH2*, *MSH6*, *MUTYH*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD51B*, *STK11*, *TP53*, and *XRCC2* were synthesized for solution hybridization enrichment. Aside from protein coding and copy number changes, IT-based sequence analysis was used to identify and prioritize pathogenic non-coding variants that occurred within sequence elements predicted to be recognized by proteins or protein complexes involved in mRNA splicing, transcription, and untranslated region (UTR) binding and structure. Mutation-associated affinity changes were computed in transcription factor (TFBSs)[2], splicing regulatory (SRBSs)[3], and RNA-binding protein (RBBSs)[4] binding sites following mutation. This approach was supplemented by *in silico* and laboratory analysis of UTR structure.

Results: Unique and divergent repetitive sequences were sequenced in 379 high-risk, patients without identified mutations in *BRCA1/2*. We identified 47,501 unique variants and we prioritized 429 variants. The methods were first applied in 7 complete genes (*ATM, BRCA1, BRCA2, CDH1, CHEK2, PALB2, TP53)* in 102 anonymized individuals (15,311 variants)[4], then validated in 287 patients in an ethics board approved study (38,372 variants)[5]. In the validation study, we prioritized variants affecting the strengths of 10 splice sites (4 natural, 6 cryptic), 148 SRBS, 36 TFBS, and 31 RBBS. Three variants were also prioritized based on their predicted effects on mRNA secondary (2°) structure, and 17 for pseudoexon activation. Additionally, 4 frameshift, 2 in-frame deletions, and 5 stop-gain mutations were identified. Multifactorial cosegregation analysis further reduced the set of candidate pathogenic variants in some families.

Conclusion: Complete gene sequence analysis followed by a unified framework can be used to interpret non-coding variants that may affect gene expression. When combined with pedigree information, complete gene sequence analysis can distill large numbers of VUS among a wide spectrum of functional mutation types to a limited set of variants for downstream functional and co-segregation analysis.

References: [1]Caminsky et al. F1000Res 3:282, 2015; [2]Lu et al. Nucleic Acids Res. doi: 10.1093/nar/gkw1036, 2016; [3]Mucaki et al. Hum. Mut. 34:557-565. 2013; [4]Mucaki et al. BMC Med Genomics. 9:19, 2016; [5]Caminsky et al. Hum. Mut. 37:640-52, 2016.

Pilot multi-gene testing in Hereditary Breast-Ovarian Cancer

Paola Carrera[1, 2] , Anna Maria Salibra [3], Giovanni Pipitone[1] , Chiara Di Resta[4], Stefania Merella[1], Sara Benedetti[1], and Maurizio Ferrari[1, 2, 3]

IRCCS San Raffaele Scientific Institute: 1 Clinical Molecular Biology Laboratory; 2 Unit of Genomics for human disease diagnosis; 3 Vita-Salute San Raffaele University, Milano, Italy

Purpose. Germinal variants in *BRCA1* and *BRCA2* account for high penetrance Hereditary Breast-Ovarian Cancer (HBOC) cases. Other genes, displaying a lower penetrance, have been associated with HBOC risk, however their incidence in the disease has not been elucidated. To define the frequency and type of variants in these genes,

we analyzed a panel of 29 genes, in a cohort of BRCA mutation-negative patients using targeted Next-Generation Sequencing.

Patients and Methodology. Among a cohort of 307 individuals with HBOC, previously analyzed for BRCA1-BRCA2 variants, 251 mutation-negative have been selected. Among these patients, a ranking was done based on prior risk to carry a germinal mutation. The top-28 at risk patients were included in the pilot study: 2 males, 26 females, mean age of onset of 44 years for breast cancer and 45 years for ovarian cancer. A validation test was conducted for the Illumina - TruSight Cancer on a Illumina - MiSeq. Validation was performed on 21 DNA samples previously analyzed by Sanger direct sequencing in genes associated with hereditary cancers. Meanwhile, also a precision study was performed, assessing the within-run repeatability and the inter-run reproducibility. Analysis of sequencing data was performed with our bioinformatic pipeline (100% sensitivity, >94% specificity on 283 variants) and based on the BWA, GATK, HaplotypeCaller, freebayes, and MiSeq Reporter tools.

Results. Panel validation provided a 543X mean coverage in the target regions (coding, included ±20 intronic). By comparing NGS results with previous Sanger data, all the 223 variant and small INDELs (up to 17 bp) were correctly called (100% sensitivity and specificity). In addition, precision of the MiSeq machine was excellent since we obtained a 100% concordance in within-run and inter-run comparisons.
Among the 28 patients included in the pilot study, we filtered 56 variants out of a total of 2721, excluding common (>5%) and benign variants. The 56 variants were distributed in high-, moderate- and low-penetrance genes: 32%, 55% and 13%, respectively. Among the 56 variants, 3 nonsense (2 novel in ATM and BRIP1; 1 already described in MLH1 and 1 already described missense in CDH1 have been classified as probably pathogenic, while the remaining (36 missense, 15 intronic and 1 in 3'UTR) have been classified as variants of uncertain significance.

Conclusion. Our pilot study results showed the presence of a probably pathogenic variant in 4 patients (14%). The rate of detection was higher than the one previously described in other cohorts; this may be due to the inclusion in the pilot study of subjects with the higher probability for a germline variant. The presence of deleterious variants in non-BRCA genes seems to be relevant in the clinic suggesting a potential clinical utility of multipanel gene testing both for affected individuals as well as for their relatives.

SESSION 4

A systematic framework for the clinical interpretation of chromosomal copy number variants

Swaroop Aradhya, Erin Rooney Riggs, Daniel Pineda Alvarez, Erica Andersen, Athena Cherry, Sibel Kantarci, Hutton Kearney, Ankita Patel, Gordana Raca, Deborah Ritter, Sarah South, Erik Thorland, Christa Lese Martin. On behalf of the ACMG/ClinGen Structural Variant Working Group.

Geisinger, Courtagen, ARUP Laboratories, Stanford University, UCLA, Mayo Clinic, Baylor College of Medicine, Children's Hospital of Los Angeles, Ancestry.com, Invitae

Analysis of germline copy number variants (CNVs), including deletions and duplications, is an established first-tier evaluation in individuals with neurodevelopmental disorders and/or multiple congenital anomalies. Even though whole genome CNV analysis has been part of routine diagnostic use for almost a decade, interpretation of the clinical significance of some CNVs remains challenging. Clinical interpretation of CNVs has also become increasingly sophisticated in

recent years due to several factors: discovery of novel genomic disorders caused by chromosomal micro-deletions or -duplications, availability of case-control studies providing a deeper view of CNVs in healthy individuals, and innovations in high-resolution microarray and sequencing technologies that allow detection of a broader spectrum of CNVs, with resolution spanning from a single gene to the full length of a chromosome. The process of CNV interpretation remains complex and relies on meaningful use of appropriate pieces of evidence to support or refute the pathogenicity of a CNV. Despite progress in recent years, discordance in CNV interpretation persists among clinical laboratories and is largely attributable to differences in selecting and weighing evidence used in classifying the clinical significance of a CNV. Existing ACMG guidelines provide a high-level conceptual framework for applying evidence in the process of interpreting constitutional CNVs in diagnostic testing. To promote consistency and transparency in CNV interpretation, the ACMG and ClinGen are currently collaborating to update the interpretation guidelines with more specific guidance on how and when to account for various types of evidence. We have devised a point-based, hierarchical scoring system to systematically evaluate relevant evidence to determine the pathogenicity of CNVs, including overlap with CNVs reported in clinically affected individuals, overlap with CNVs reported in unaffected individuals, case-control studies, presence of known dosage-sensitive genes, case reports with segregation data, de novo occurrence of CNVs, and the number of protein-coding genes included in the CNV. We describe the rationale behind this new framework and provide a detailed rubric for interpreting chromosomal deletions as the first step in our process. We also present data from testing this framework with a set of CNVs among clinical cytogeneticists and laboratories. This collaborative group will further expand the interpretation rubric to include chromosomal duplications, CNVs that do not contain known protein coding genes but may involve important regulatory elements, and CNVs that affect single genes or involve only part of a gene. The new CNV interpretation framework will be tested with a broad group of cytogeneticists to identity nuances and fine-tune the relevant guidance. The updated guidelines will apply to analysis performed by any method capable of defining chromosomal boundaries of copy number events, even from routine whole exome or genome sequencing. This work is expected to have broad impact in the clinical community by providing a robust system to support the consistent interpretation of CNVs.

Challenges in Variant Interpretation: How to minimize inter- and intra-laboratory inconsistencies

A. Laner, A. Benet-Pages, E. Holinski-Feder

MGZ – Medizinisch Genetisches Zentrum, Bayerstr. 3-5, 80335 Munich, Germany

Establishing guidelines always means convincing people to give up their traditions and customs and to accept the concepts and ideas of others. This applies to all kinds of human interaction in general and to geneticists in particular. Standardization in gene nomenclature, variant nomenclature, and variant classification, although now generally accepted, did not come naturally or without resistance.

In the case of variant classification guidelines, a major argument against the uniform use of the ACMG guidelines (1) was poor inter-laboratory consistency, as documented in early studies (2) (3). The inconsistencies arose for several reasons, e.g., the selected variants were challenging, pushing criteria to the limit; lack of previous training in how to deal with conflicting lines of evidence or examples and when to ignore certain contradictory evidence; some ACMG criteria are somewhat “blurred” and leave room for misinterpretation, etc.

However, these initial comparisons between laboratories were positive, at least in the sense that the guidelines were useful in providing a common framework for facilitating the resolution of differences between the participating laboratories.

In considering these results we asked the question of whether 10 molecular biologists and/or assessors in our laboratory showed consistency in classification when using the ACMG guidelines, and, if not, what specific problems in variant interpretation could be observed. For this purpose we set up a proficiency test in which each participant had to classify 20 variants selected to cover a broad range of ACMG qualifiers. In evaluating the test results, we identified some key parameters that accounted for most of the inconsistencies. These parameters were 1) identifying and extracting the relevant piece of evidence from the literature, 2) weighting specific evidence given in publications, and 3) identifying and overriding obviously poor data or evidence for an otherwise consistent classification.

We answer these problems with regular meetings, in which we discussed some of these challenging variants, and with workshops, where we train our staff to operate in harmony. Furthermore, we observed that the display of relevant variant data on our interpretation software is a critical part of the process as well, and we will continue to improve on this. Interestingly, no specific error-prone ACMG criteria were identified in this setting.

(1) Richards, et al., Standards and guidelines for the interpretation of sequence variants; *Genet Med.,* March 2015

(2) Amendola, et al., Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium; *Am J Hum Genet*.; June 2016

(3) Maxwell, et al., Evaluation of ACMG Guideline-Based Variant Classification of Cancer Susceptibility and Non-Cancer-Associated Genes in Families Affected by Breast Cancer; *Am J Hum Genet.,* June 2016

From a list of variants to a diagnostic report: extracting clinically relevant information

María-Jesús Sobrido & Beatriz Quintáns

Neurogenetics Research Group, Instituto de Investigación Sanitaria, Santiago de Compostela & CIBERER.

Whole exome sequencing (WES) is already being used in the clinical setting, quickly evolving into standard of care. WES may be a first choice for rare syndromes, for disorders without a clear-cut diagnostic suspicion, as well as situations where treatment outcome may radically depend on early diagnosis. Besides technical challenges of sequencing and computational processes, the main barriers to a widespread implementation of NGS in diagnostic laboratories come from a limited ability to provide adequate medical interpretation and reporting for many of the identified variants. Out of the long variant lists obtained through NGS, only those that are *clinically relevant* (or, equally ambiguous, *actionable*) should probably end up in the final report. But who and how decides what is *clinically relevant/actionable* in a given case? Ideally, the reported result would only contain one or a few genetic variants that unequivocally explain the patient´s symptoms. However, as the capacity of genomic technologies increases, so will also the observation of unanticipated, potentially relevant variants – dubbed incidental or secondary findings - as well as a plethora of variants of unknown significance, that may add a heavy burden on genetic counselling. There is no consensus on how to define, analyse, and report such variants to patients and physicians, who need fast and cost-effective reports with useful clinical insights.

Published studies on WES for rare disorders yield diagnostic rates ~30-60%. One fundamental pitfall, however, to establish the real diagnostic rate is that for most cases the causality of the variants has not been proven. Our judgement may be misguided by: i) available information in the (imperfect) repositories, ii) what we know and don´t know about our genome, iii) lack of distinction between statistical significance and clinical significance, iv) lack of distinction between biological evidence and medical evidence, v) fragmented interpretation of the genomic findings (one variant at a time), vi) our own previous experience. Interpretation is particularly challenging for diseases with only a few cases reported. For most genetic disorders, the pathophysiological mechanisms are unknown, lending evaluation of a variant´s downstream consequences extremely difficult. It might eventually turn out that many patients diagnosed in these initial years of clinical NGS received incorrect interpretations of their results. It is hardly conceivable that a unique set of rules for assessment of pathogenicity – let alone causality – will be applicable to every disease, gene and variant type. Experimental methods to evaluate functional effects of variants are not amenable to systematization for many genes and diseases. Therefore, the process of variant interpretation will generally require the collaborative work of different experts, including biologists, bioinformaticians and clinicians.

While the technical aspects of variant detection keep improving, there is an urgent need for studies on clinical interpretation and delivery. These studies may include enquiries on what doctors think NGS reports should contain. Do they expect that the laboratory provides interpretation, or should the genetic variants just be described in the report and interpretation left to clinicians? Multidisciplinary working groups should generate (and periodically revise) recommendations on interpretation and reporting genetic findings in their field of expertise, providing guidance on issues like how to weigh the evidence and classify variants, reliable databases in that specialty, and what to include in the report: primary findings, secondary findings, variants of unknown significance, literature information, level of evidence of pathogenicity for each variant, follow-up recommendations, whether results will be re-evaluated on regular basis, etc. These working groups could also carry out quality control schemes and benchmark clinical interpretation and reporting processes of diagnostic laboratories.

Accessing the full size-spectrum of human genetic variation using PacBio long-read SMRT sequencing on the Sequel System

Luke Hickey

Senior Director of Human Biomedical Sciences, PacBio

Using PacBio long-read SMRT sequencing methods, it is now possible to sequence and *de novo* assemble an individual human genome to construct a reference-quality diploid assembly that accurately represents and phases both maternal and paternal haplotypes [1]. Global initiatives are currently underway to apply these *de novo* assembly methods to individuals representing global populations, thereby extending the diversity of the current Human reference genome [2, 3, 4, 5]. Analyses comparing these population-specific *de novo* genomes to the standard admixed human reference (Hg19, GRCh38) provides a more comprehensive view of individual human genetic diversity, identifying upwards of ~16 Mbp of unique sequence in each individual genome that is not currently represented in the reference [3, 4, 5, 8].

In addition to identifying novel genomic sequence, these efforts have demonstrated that any individual diploid human genome contains upwards of ~20,000 unique structural variants (defined as >50 bp in size), and another ~400,000 indel variants (ranging in

size from 1 bp to 49 bp) [3, 4, 6, 7, 8]. Over 80% of these variants [6] are not currently accessible using short-read WGS methods, due to coverage and mapping bias inherent in SBS technology-dependent approaches.

New PacBio sequencing methods including; low-fold long-read WGS, targeted long-read sequencing, and visualization tools (IGV 3.0 [10], Ribbon [11]), have made it feasible and efficient to access and study the full size-spectrum of human genetic variants in-phase, across individual human genomes, with an economical study-design.

By applying these new PacBio long-read sequencing methods to Precision Medicine and Human Population studies, human biomedical research scientists have recently reported discovery of novel causative genetic variants in both rare [9] and common disease studies [12], demonstrating the value of a long-read approach.

I will describe these novel long-read sequencing methods and research discoveries, including a new Human genome sequencing dataset generated on the Sequel System with our latest chemistry and library preparation methods.

REFERENCES:

[1] Chin CS, et al. (2016), *Phased diploid genome assembly with single-molecule real-time sequencing*. Nature Methods 13, 1050–1054.

[2] Graves-Lindsay T, et al. (2016), *McDonnel Genome Institute Reference Genome Improvement Project*, http://genome.wustl.edu/projects/detail/reference-genomes-improvement/

[3] Shi L., et. al. (2016) *Long-read sequencing and de novo assembly of a Chinese genome*, Nature Communications, 7:12065.

[4] Seo JS, et al. (2016), *De novo assembly and phasing of a Korean human genome.* Nature, 538(7624):243-7.

[5] Tohoku Medical Megabank Organization [ToMMo] (2016), Japanese Reference Genome (JRGv1), https://jrg.megabank.tohoku.ac.jp/en/

[6] Huddleston J, et al. (2016), *Discovery and genotyping of structural variation from long-read haploid genome sequence data,* Genome Res. gr.214007.116.

[7] Chaisson MJ, et al. (2015), *Resolving the complexity of the human genome using single-molecule sequencing*, Nature, 517(7536): 608-11.

[8] Pendleton M, et al. (2015), *Assembly and Diploid Architecture of an Individual Human Genome via Single Molecule Technologies*, Nature Methods 12, 780-786.

[9] Merker J, et al. (2016), Long-read whole genome sequencing identifies causal structural variation in a Mendelian disease, bioRxiv 090985.

[10] Robinson JT, Integrative Genomics Viewer (IGV), Using the Integrative Genomics Viewer to visualize PacBio long-read SMRT Sequencing data.

[11] Nattestadt M, et al. (2016), *Ribbon: Visualizing complex genome alignments and structural variation,* bioRxiv 082123

[12] Lutz M, et al. (2016), *Identification and characterization of informative genetic structural variants for neurodegenerative diseases*. PacBio ASHG 2016 Workshop Live

SESSION 5

Challenges with the compilation and naming of new variation revealed by massively parallel sequencing of forensic markers

Chris Phillips

Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Galicia, Spain

At the start of 2016, the ISFG DNA commission on minimal nomenclature requirements was formed. This group was established to oversee the annotation of new variation detected by use of massively parallel sequencing (MPS) to genotype the core forensic markers used across the globe. The commission published a set of guidelines [1] that addressed: sequence alignment based on an agreed single sequence strand aligned to the GRCh37 and GRCh38 human genome builds; variant annotation from the compilation of known polymorphisms plus novel sequence variation in the full sequence string obtained from MPS analyses; and a minimal nomenclature system covering the whole sequence string and anticipating hitherto unrecorded variation.

Here we report updated frameworks for sequence alignment and variant annotation, with emphasis on the major challenges emerging from using MPS to type forensic markers. These challenges include: imprecise mapping of Indels in STR repeat regions; characterizing common STR repeat region sequence that is not found in the reference sequence; annotating private or very low frequency nucleotide substitutions; tracking variation that creates or disrupts the counted repeat units of STRs (i.e. alleles which populate the national DNA databases); and simple, but descriptive nomenclature rules for microhaplotypes.

[1] Parson W, et al, Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements, Forensic Sci. Int. Genet. (2016) 22:54-63.

Naming Genetic Variation in Forensic Science: Alignment and Nomenclature of Next Generation Sequence Alleles

Walther Parson[1,2]

[1] *Institute of Legal Medicine, Medical University of Innsbruck, Austria*
[2] *Forensic Science Program, The Pennsylvania State University, PA, USA*

Unambiguous and consistent naming of genetic information is crucial in forensic science and forensic routine practice. DNA profiles derived from genotypes and haplotypes serve as a means of communication in forensic reports and form the basis for quantifying genetic evidence for court-related purposes. The forensic community is establishing genetic information from known and unknown samples by capturing either fragment size-based variation, e.g. Short Tandem Repeat (STR) analysis by electrophoresis, or sequence-based differences, e.g. Single Nucleotide Polymorphisms targeted in nuclear and mitochondrial DNA. Nomenclature rules were established for the first type of variant that routinely translate repeat numbers into allele categories, while the second type are often reported relative to selected reference genomes. The transition from electrophoresis to massively parallel sequencing (MPS) techniques poses a challenge to the community, as the nomenclature of the newly accessible information from the full nucleotide sequences of STR loci needs to remain backwardly compatible to the millions of STR genotypes generated by electrophoresis that form a huge and important body of

information in national DNA databases for practical and forensic intelligence purposes. New alignment and nomenclature rules are being developed and facilitated with customized software in an attempt to comply with these requirements. The challenge of establishing new nomenclature systems that embrace genetic data from different technological developments lies in the adoption and modification of earlier conventions that were based on the limitations of electrophoretic analytical methods in use for almost twenty years. For STR loci, for example, this includes the choice of the sequence strand relative to which repeat motifs have been previously aligned and reported. MPS applied to the field of mitochondrial DNA, offers numerous analytical advantages, but can show as yet unobserved limitations in the quantitative reporting of homopolymeric regions, which requires adoption of new interpretation rules with compromised accuracy compared to earlier data.
The presentation provides a guide through the achievements established in the early days of forensic DNA fingerprinting and contrasts these current developments for the implementation of nomenclature systems for that capture a broader range of genetic variation derived from modern sequencing techniques.

BaseSpace Variant Interpreter: A new platform to improve the speed of genomic interpretation and facilitate collaborative knowledge sharing

Jennifer Harrow (on behalf of BSVI team)

Bioinformatics Dept, Illumina Inc, 5200 Illumina Way, San Diego, 92122 United States.

Significant reductions in the of cost genomic analysis are making the use of clinical whole genome sequencing a valuable tool for oncologists and rare disease specialists alike. The rate of accurate analysis and interpretation of the millions of variants coming from such clinical initiatives, however, remain a challenge for rapid sample to answer paradigms. BaseSpace Variant Interpreter (BSVI) is a web-based interpretation and collaborative knowledge sharing tool to expedite annotation, filtering and interpretation of genomic data from patients with rare disease or somatic cancer. BSVI supports different interpretation workflows and aggregates annotations from a broad range of public and private knowledge bases to help triage variants. While BSVI has been optimized for whole genome datasets (gVCFs), it can operate on a range of sequencing scales. Rapid data ingestion (3-5 minutes) and subsequent filtering operations (seconds) make BSVI an effective and efficient analysis tool. Variants can be filtered by data quality, predicted consequence, population frequencies, variant impact and inheritance with support for custom gene or region lists for bespoke filtering. Results can be ordered by pathogenicity or actionability for further review and validation. Filters can be optimized for specific phenotypes or tumor types using ontology terms from HPO or SNOMED.

BSVI is designed from the ground up to support data and knowledge sharing. The concept of workgroups and domains permits users to collaboratively access and share cases with geographically distinct users and groups. BSVI permits users to add new evidence-based curations; knowledge that is shared within workgroups and helps prioritize future occurrences of the variant in other cases. Detailed audit logs in BSVI track changes to annotations and monitor how the data has been modified and assess the validity of the evidence used. BSVI is being developed in collaboration with Genomics England. We intend to showcase BSVI and describe how we are beginning to capture bespoke data from the labs and integrate into a crowd-sourced knowledgebase available to the Genomic Medicine Centres (GMCs) of the NHS.

Safe variant annotation sharing across laboratories

Beat Wolf 1,2, Pierre Kuonen 1, David Atlan 3, Jonathan Stoppani 1, Davide Mazzoleni 1, Thomas Dandekar 2

1HES-SO Fribourg, Switzerland
2University of Würzburg, Germany
3Phenosystems SA, Belgium

Purpose
Annotating genetic variants heavily benefits from the usage of databases that already document the concerned variants. Due to the massive amount of possible variants and their complex effects, data sharing is essential to improve understanding of their consequences. Unfortunately many variants are only documented in private databases, either because of legal restrictions or because of the lack of time to publish the findings on public databases.
To solve this problem, we want to extract the information from where it already exists, which are the NGS data tools used in various laboratories. They often already integrate an internal variant database. We propose to augment those tools in a way to make it possible to share the variant annotations anonymously and securely among other users of the software, while giving the users full control on how to share their data.

Methodology
For our first prototype we choose to extend the graphical NGS data analysis software GensearchNGS, developed by Phenosystems SA, with the ability to share variant annotations across multiple laboratories. Our approach is based on an original concept we call Trusted-Friend-Computing (TFC). The basic idea of THC is the ability to share resources through a network of trusted friends. To implement this concept in GensearchNGS, we extend the open source POP-Java programming language with the required functionalities. POP-Java is a Java language extension that dramatically eases development of distributed Java applications. The TFC concept also allows us to explore the possibility to share computing resources across multiple laboratories, but, in the current project, we will focus on variant annotation sharing.

Results
We present the prototype of the proposed functionality in GensearchNGS as well as its open source implementation in POP-Java. We present the current state of the prototype as well as the underlying design principles.

Conclusion
Sharing variant annotations is essential to improve the quality NGS data analysis. We acknowledge the technical and legal restrictions which hinder wide spread sharing, and propose thus a more controlled way to share variant annotations among a group of trusted laboratories. We hope that our work can serve as a stepping stone towards a more open variant data sharing culture.

Variant detection and the challenges beyond: what is needed to implement genetic information for clinical use?

Daniela Steinberger

Human Geneticist, Medical Director, bio.logis Center for Humangenetics, Frankfurt am Main, Germany

Methods and instrumentation to produce genotyping data are improved at an increasingly fast pace. As a consequence, the availability of diagnostic tests based on high-throughput analyses up to whole genome sequencing will be gradually established as a routine for theoretically every citizen. One of the biggest challenges will be the translation of the large amounts of genetic data into useful and comprehensible medical information and recommendations. The presentation gives an overview of the status quo for conventional genetic diagnostic

reporting and communication processes as well as the needs that are arising with the advent of the new powerful genotyping methods. Possible solutions to overcome these challenges are discussed in order to enable the realization of a medical benefit based on the large amounts of personal genetic data.

SESSION 6

Challenges in bioinformatics for genetic diagnosis

Joaquin Dopazo

Clinical Bioinformatics Research Area, Fundacion Progreso y Salud, Sevilla, Spain

Precision medicine aims to find better ways of defining diseases by gradually substituting conventional clinical or pathological diagnostic criteria with state-of-the-art genomic profiling methodologies. More precise diagnostic of diseases, based on the description of their molecular mechanisms, will foster the advent of innovative diagnostic, prognostic, and therapeutic strategies precisely tailored to each patient's requirements. To achieve this, different technical challenges related with genomic data management, integration and storage must be faced. Furthermore, in addition to technical challenges, a deeper, systems-based understanding of these disease drivers is required in order to find more precise biomarkers.

Here I will show efficient solutions to address challenges associated to genomic data management in the context of conventional pipelines of variant prioritization. Moreover, I will discuss how primary gene variation data, alone or integrated with other omic data, such as gene expression, can be transformed into high-precision mechanism-based biomarkers containing higher-level information on disease mechanisms and drug modes of action (MoA) using simple computational models.

A community-developed data model for representing sequence variation

Reece K. Hart* (Invitae, GA4GH)*; Gil Alterovitz (Harvard Medical School, FHIR Genomics); Larry J. Babb (Sunquest, ClinGen); Karen Eilbeck (University of Utah, Sequence Ontology); Robert R. Freimuth (Mayo Clinic, HL7 Clinical Genomics Working Group, ClinGen); Sarah E. Hunt (EBI, Ensembl); David Kreda (Harvard Medical School, HL7 Clinical Genomics Working Group); Jennifer Lee (NCBI, ClinVar); Peter N. Robinson (Jackson Laboratory, Human Phenotype Ontology); Shawn Rynearson (University of Utah, Sequence Ontology); David Haussler (UC Santa Cruz, GA4GH); Heidi Rehm (Harvard Medical School, Broad Institute, ClinGen); Peter Goodhand (Ontario Institute for Cancer Research, GA4GH)

* reecehart@gmail.com

Maximizing the personal, research, and clinical value of genomic information will require that clinicians, researchers, and testing laboratories exchange genetic variation data reliably. The Variation Modeling Collaboration (VMC) is a partnership with representatives from the ClinGen, NCBI ClinVar, FHIR Genomics, GA4GH, HGVS, and HL7 ClinGen Working Group communities. Our goal is to propose and encourage the adoption of a data model with clear semantics for the computational representation of variation on any sequence type. The VMC model will be interoperable with other formats and enable the reliable exchange of sequence variation among data providers and consumers.

The VMC draft specification precisely defines three essential genetic states — allele, haplotype, and genotype — and provides conceptual data models for each. Allele represents a contiguous sequence (reference or alternate) with respect to a reference sequence. Haplotype represents alleles occurring on the same molecular sequence (that is, "in phase" or "in cis"). Genotype

represents a list of haplotypes, thereby providing a representation of Alleles at any number of sites and observed with various specified ploidy due to chromosomal or segmental aberrations, or clonal mixtures such as heteroplasmy, mosaicism, and chimerism. Alleles, haplotypes, and genotypes may be defined on DNA, RNA, and protein sequences.

In addition to its core models, the VMC draft specification includes additional guidance to standardize data exchange in "real world" uses. As such, it includes models to represent human-readable identifiers for instances of genetic variation, such as references to ClinVar variants or names for haplotypes. It also provides models to represent multiple notions of relationships between concepts, such as alleles related by normalization or translation. It also specifies and recommends an algorithm for constructing distributed and globally-unique identifiers for genetic states. Importantly, this algorithm enables two parties to share variation data without prior agreement about an assigned identifier. Finally, the draft specification provides technical guidance for transmitting models using common serialization technologies.

Meiosis, recombination, and the origin of a species

Peter Donnelly

The Wellcome Trust Centre for Human Genetics, University of Oxford

Abstract: In many mammals, including humans and mice, meiotic recombination events occur in a set of small genomic regions called recombination hotspots. The gene *PRDM9* is now known to be responsible for positioning the double-strand breaks which initiate recombination, via its zinc-finger domain which binds to particular motifs in genomic DNA. There is extensive variation in the zinc finger array between, and often within, species, and these differences in PRDM9 binding result in quite different sets of hotspot locations. Completely separately, *Prdm9* has also been shown to play an important role in speciation. In fact, it was the first speciation gene identified in vertebrates, and remains the only such gene known in mammals, being responsible for hybrid infertility in the offspring of crosses between certain mouse subspecies. We re-engineered the zinc finger domain of *Prdm9* in the C57BL/6 laboratory mouse, replacing the wild-type domain with that from the human reference allele. This change repositions hotspots and completely restores fertility in the hybrids**.** Various analyses allow us to uncover the molecular mechanism driving the hybrid sterility, and in so doing to establish an important role for PRDM9 in meiosis downstream of its role in positioning double-strand breaks.

Raising the bar in NGS diagnostics : Challenges in Variants Identification

Zélie Dubreucq

Subject Matter Expert, Sophia Genetics

Today Sophia Genetics is the global leader in Data-Driven Medicine. SOPHiA™, the collective artificial intelligence for clinical genomics processes and analyzes huge amount of data every day, understanding all the challenges and the bias associated to NGS technologies. Sufficient coverage of target region is a well-known prerequisite of accurate variant detection. However, there are other issues that are less obvious but important to ensure correct variant identification. For example, variants exposed to the end of reads, proper trimming of the primer sequences, special care of repetitive regions, etc. Understanding the limitation of the sequencing platform, the chemistry used for enrichment, the sequence context (gene panels) and the sample type is the key to ensure an accurate analysis workflows.

Questions such as how many samples should be multiplexed in a single run, how long will the indel be detected, what will be the limitation of the analysis are often asked by the routine diagnostic labs. Here I introduce the efforts that we made to tackle those different problems normally encountered in routine genetic testing using NGS.

SESSION 7

Strategies for assembling high quality genome sequences

Ivo Glynne Gut

Centro Nacional de Analisis Genomico, Center for Genomic Regulation, Barcelona, Spain

At the CNAG we have carried out *de novo* genome assembly and annotation for many different species (*e.g.* Iberian lynx, turbot, cedar aphid, almond, wasp, olive). The objective of each *de novo* assembly is to provide high contiguity and scaffolding with correct order and orientation of contigs so that downstream annotation has the best possible chance to recover the genes and gene structures correctly. Over the years we have refined the sequencing strategies that we apply to optimize contiguity and scaffolding while reducing the overall cost of a *de novo* assemble. Our basic strategies use paired-end whole genome shotgun sequencing together with matepair analysis, fosmid pool shotgun sequencing and fosmid end sequencing in pools that are all run on Illumina sequencing systems. Computational assembly strategies start with de novo assembly of fosmid pools, followed by error correction with the whole genome shotgun data. Scaffolding is refined with matepair and fosmid end sequences. Throughout we apply a contig breaking and re-assembly strategy. We are constantly exploring datatypes from different types of sequencers to optimize the quality of an assembly and the cost of building a new genome sequence. In particular these are the inclusion of fosmid pool sequencing using Pacific Biosciences and Oxford Nanopore Technologies MinIon systems and optical mapping strategies with the Iris system from BioNanoGenomics. With the improvements of the quantity, read length and quality of sequences obtained from the ONT MinIon V9.4 we attempted to directly sequence several gigabase-size genomes. We achieved *de novo* assembly with contig N50 greater than 1 Mbase off the bat. In this presentation we will discuss the strategies and pipelines that we have developed to achieve high quality reference genomes and annotations for the different *de novo* projects carried out at the CNAG.

Latest applications of innovative technologies in research and diagnostics

Johan T. den Dunnen

Leiden Genome Technology Center (LGTC), Human Genetics & Clinical Genetics, Leiden University Medical Center, Leiden, Nederland

Research and diagnostics are often technology driven; new technologies give new options to perform analysis that were initially not possible or not affordable. Our institute hosts a facility for genomics and transcriptomics technology driving the development of new approaches and sparking collaborations with scientist in many different disciplines. A snapshot of these developments will be presented, giving specific attention to the underlying methodology used.

Current interests, besides the general application of next-generation sequencing technology, focus especially on single molecule sequencing and single cell ananlysis. Standard next generation sequencing (Illumina), combined with single molecule sequencing (Pacific Biosciences), is used to determine the complete genome sequence of genomes from bacteria to snakes

with a focus on determining the structure of complex regions with repetitive DNA sequences. Long-read single molecule sequencing is used to study complex regions in the human genome, incl. genes like PKD1, CYP2D6 and the 3.3kb D4Z4 repeat (FSHD), as well as to perform full length mRNA sequencing resolving RNA structure and potential long-distance RNA splicing dependencies. Single cell RNA sequencing is used to study the composition of cell cultures/ tissues and to identify rare cell fractions (<1%) expressing specific genes of interest. 10x genomics technology, combined with whole exome sequencing, is used for haplotyping and to increase the sensitivity to detect structural variants. Ribosome profiling was developed to study actively translated RNA, protein translation and uORF sequences.

Variant detection, interpretation and reporting using an end-to-end analysis pipeline

Ruth Burton[1], Shyamal Dilhan Weeraratne[2] and Julie Deschênes[2]

[1]QIAGEN Advanced Genomics, Manchester, UK, [2]QIAGEN, Waltham, USA

Historically, detection, interpretation and reporting of variants has been a labour intensive and time-consuming process. A range of different software applications have been required to detect and filter variants and information from a variety of sources has been needed to accurately classify each variant. This process can be further complicated if there is a need to provide up to date information on relevant treatments and clinical trials. The final step of collating and summarising findings in a clinical report completes the process.

Using datasets from The Cancer Genome Atlas (TCGA) we present an end-to-end workflow starting with raw, unaligned reads from whole exome sequencing (WES) and ending with a comprehensive report describing clinically relevant variants. The first stage of the data analysis pipeline - to align the next generation sequencing (NGS) reads and detect variants was performed using Biomedical Genomics Workbench (BxWB). The list of variants was then refined using Ingenuity Variant Analysis (IVA). This filtering step, critical for WES, formed an integral, seamless part of the analysis pipeline in BxWB. The final steps of interpretation and report generation were performed by transferring the filtered variant list from IVA into QIAGEN Clinical Insight (QCI™) Interpret software and the automatic application of the American College of Medical Genetics (ACMG) Guidelines [1].

The combination of these three powerful software packages in a single pipeline were shown to reliably identify and classify clinically relevant variants. We were able to detect and classify disease causing variants in two colorectal cancer datasets including the KRAS variant p.G12R and the TP53 variant p.R273H. In addition using four breast cancer datasets we detected disease causing variants in BRCA2, PIK2CA, TP53 and MAP2K1. We were able to contextually associate these clinically significant variants with relevant treatment information and clinical trial recommendations to support patient care.

1. Journal article - Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody W, Hegde M, Lyon E, Spector E, VoelkerdingK, Rehm H (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine 17:405-423 doi: 10.1038/gim.2015.30

Rapid Genome Wide Mapping at the Single Molecule Level Using Nanochannel Arrays for Structural Variation Analysis and de novo Assembly

Dr. Sven Bocklandt

Senior Application Specialist, Bionano Genomics

The promise of genomics in relation to human disease is being held back by the inability to resolve large structural variations. Existing technologies, including next-generation sequencing (NGS), diagnose less than 50% of patients with genetic disorders. Large structural variants such as deletions, duplications, inversions and translocations are extensively present, and many are known to affect biological functions and cause disease, including cancers and developmental disorders. Bionano genome mapping offers unmatched structural variation discovery, making Saphyr essential to human genome and translational research.

With the new Saphyr, genome maps are generated in high speed & throughput from massively parallel single-molecule visualization of extremely long DNA molecules without amplification, providing long-range contiguity critical for *de novo* sequence scaffolding and analysis of structural variation in complex genomes. Large structural variants and repeats can be visualized and measured directly within long range genome information for comprehensive analysis.

SESSION 8

Advancing Drug Discovery - Translating Small & Big Data Into Insight

Gunaretnam (Guna) Rajagopal, PhD

VP Global Head – Computational Sciences, Discovery Sciences, Janssen Research & Development (Johnson & Johnson), Spring House, Pennsylvania, USA

A central challenge on drug discovery and development involves our ability to capture, interrogate and interpret data from diverse sources from the bench and the clinic and generated within and outside Janssen. For example, ongoing initiatives in precision oncology aims to combine comprehensive data collected over time to enable disease understanding to advance disease prevention, interception and cure. To this end, considerable resources are being expanded to harness the potential value of data derived from lab-based assays, electronic health records, ‘omics technologies, imaging, and mobile health etc. We will review our ongoing efforts to translate data into insight, through selected examples, of how we are exploiting computational analytics and experimental capabilities in partnership with scientists within and outside of Janssen to address and support the discovery of safe and effective therapies that will benefit patients.

INNOPHARMA: the Pharmacogenomics platform of the University of Santiago de Compostela

Eduardo Domínguez, José Brea and María Isabel Loza*

Innopharma pharmacogenomics platform. University of Santiago de Compostela. Spain

The new model of "open innovation" is emerging in the pharmaceutical sector. Thus, opportunities for pre-competitive research are increasing and the chance to develop open innovation models is becoming apparent.

In Galicia (Spain), there is a gap between the basic research on new therapeutic mechanisms and the industrial application of the findings of such research. Thus, extremely interesting results coming from genomic research identifying promising therapeutic targets did not advance up to novel drugs due to this gap. INNOPHARMA pharmacogenomics platform was created to leap over this gap by undertaking projects that will add value to scientific findings and transfer the corresponding knowledge to drug discovery programs. INNOPHARMA was funded by the EU Technology Fund, via the Spanish Ministry of Economy and Competitiveness and the Galician Innovation Agency.

The platform was based on three main pillars:

- A high throughput pharmacogenomics platform of innovative technologies applied to assay panels in early drug discovery
- Constant development of an outstanding chemical library of compounds with chemical and biological diversity, including repurposing
- A pipeline of collaborative programs with industrial standards, allowing the establishment of strategic partnerships in a context of open innovation.

In its first call for research projects, INNOPHARMA received 128 proposals, being ten of them selected for further development by means of carrying out the studies required to exploit the results of the public groups and develop joint projects in early drug discovery in different therapeutic areas. One of these collaborative projects entered into the GSK Fast Track to Innovation program for further development. INNOPHARMA platform is open to receive proposals about those genomic findings that could be relevant for developing novel human therapies.

Advances in Personalised Medicine – Liquid Biopsy

Dr. Alexander Sartori

Agena Bioscience GmbH, Germany

The analysis of Circulating tumor cells (CTC) and circulating tumor DNA DNA (ctDNA) is becoming increasingly important as an alternative source for detection of actionable mutations in cancer due to the tumor heterogeneity as well as the practical limitations and negative implications for the patient that go along with invasive tissue biopsies. Recent technological developments enable the use of CTCs and ctDNA as cancer biomarkers as a non-invasive complement and potentially even substitute for tissue biopsies. We will discuss the current research applications of ctDNA in cancer detection, and how the highly sensitive UltraSEEK™ Technology with the MassARRAY™ System can enable fast, easy-to-perform and cost-effective liquid biopsy testing.

SESSION 9

PacBio long read sequencing for improved resolution of complex genomic variation

Henk PJ Buermans1, Rolf H Vossen1, Yahya Anvar1,2 ,Tahar van der Straaten2 Daniel M. Borràs3, Stefan J White1 and Johan T den Dunnen1,4

1: Leiden Genome Technology Center, Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
2: Department of Clinical Pharmacy & Toxicology, Leiden University Medical Center, Leiden, The Netherlands
3: Genomescan B.V, Leiden, The Netherlands
4: Department of Human & Clinical Genetics, Leiden University Medical Center, The Netherlands

The emergence of genomic technologies has led to unprecedented insights into the role of genetic variants in human biology. Although variant detection by classical Sanger sequencing is gradually being replaced by next generation sequencing (NGS) technologies, it will not be able to resolve all genomic variation. A major limitation of both methods is their short read length (<1,000 nucleotides), making it impossible to resolve certain rearrangements and/or elucidate regions with high similarity or low complexity sequences. Genomic variation within these regions are often not detected or inaccurate due to alignment ambiguity and/or insufficient read information. PacBio single molecule long-read sequencing can overcome these limitations and tackle some of the remaining diagnostic problems. The current system yields average read lengths of 15 kb with context-free error profile, allowing for high-quality consensus reads of >QV50 (one error in 105bases) to be generated from the relatively high-error single-pass data. We illustrate the advantages of long read sequencing for several use cases.

The Cytochrome P450 2D6 (CYP2D6) gene is among the most important genes involved in drug metabolism. Specific variants in the CYP2D6 gene determine the enzyme's amount and activity. The CYP2D6 locus is specifically enriched by long range PCR (LR-PCR) as a 6.6kb fragment using primers that exclude the CYP2D7 and CYP2D8 pseudogenes. We obtained high-quality, full-length, phased CYP2D6 sequences, enabling accurate variant calling and resolution of the separate haplotype sequences of the entire gene-locus including exonic, intronic and up and downstream regions. Diplotypes used for CYP2D6 protein activity prediction were confirmed for 24 of 25 samples, including gene duplications, and variants were detected that had not previously been associated with specific haplotypes.

Genetic diagnosis of autosomal dominant polycystic kidney disease (ADPKD) is challenged by gene allelic heterogeneity, high GC content, and homology of the PKD1 gene with six pseudogenes. PKD1 was specifically enriched by 5 partially overlapping LR-PCR fragments. Primers were chosen such that they effectively excluded off-target signals from the pseudogenes. Comparison of PacBio long read data with Sanger sequencing showed a high concordance between the two methods.

Conventional mitochondrial-DNA (MT DNA) sequencing approaches use Sanger sequencing of 20-40 partially overlapping PCR fragments per individual, which is a time- and resource-consuming process. We have setup a high-throughput, accurate, fast, and cost-effective human MT DNA sequencing approach using PacBio sequencing. In this setup we first generate LR-PCR products for two partially overlapping 7.7 and 9.2 kb MT DNA-specific amplicons, add sample-specific barcodes, and sequence these on PacBio to obtain full-length MT DNA sequences for genotyping/haplotyping purposes.

In conclusion, long sequence reads are pivotal to accurately identify and exclude off-target signals from homologous sequences in the genome of interest, such as pseudogenes. We conclude PacBio-sequencing is a powerful diagnostic tool facilitating analysis of complex genomic loci and provides a robust framework for improved disease diagnostics.

Tracing mitochondrial mutations in 3D in primary tumor, lymph node and liver metastasis

Paulo Refinetti[1], Stephan Morgenthaler[1], William Thilly[2] and Per O. Ekstrom[3]

[1]Ecole Polytechnique Federale de Lausanne, Switzerland; [2]Massachusetts Institute of Technology, USA, [3]The Norwegian Radium Hospital, Norway

Background

Lineage tracing provides a powerful means of understanding tissue development, and tumor heterogeneity [1,2]. Tumors start as a single clone to become a diversified mass[3,4]. Recent studies have shown adjacent areas to differ in gene expression, morphology or metabolism[5].

Methodology

Continuously serial sectioned tumor tissue is subjected to laser capture micro-dissection and mutation analysis. Mutant fractions is determined by cycling temperature capillary electrophoresis in total of 9500 laser capture micro-dissected samples, taken from 8 different tissue samples. Iso-surfaces are used to reconstruct a 3D representation of the mutated areas within the samples volume (1,25mm^3). User interactive 3D presentation of the data can be found at the following web site; http://www.ous-research.no/hovig/?k=hovig%2FEkstrom&aid=16790

Results

The three-dimensional lineage tracing in human tumor volumes and metastasis revealed non-uniform patterns of cell growth. The lineages followed are highly diffused through the volume analyzed, occupying only a fraction of the volume. Our observations could be explained by heterogeneous tumors or the infiltration and remodeling of host tissue by tumor cells. Following a cell clone from the primary tumor to a metastasis showed similar results. If a single cell starts a metastasis it must grow by infiltrating and remodeling the local host tissue to produce the observed results, otherwise metastasis must have a cell cluster as origin[6].

Conclusion

The method proposed can be used to study other human tissue both pathological and healthy. The high spatial resolution combined with the possibility of three-dimensional analysis will permit the observation of previously unsuspected patterns. These patterns in turn could change the general understanding of stem cell behavior in both healthy organs and tumor tissue.

Refferences

1.Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501,** 328–337 (2013).

2.Kretzschmar, K. & Watt, F. M. Lineage Tracing. *Cell* **148,** 33–45 (2012).

3.Fialkow, P. J. Clonal Origin of Human Tumors. *Annu. Rev. Med.* **30,** 135–143 (1979).

4.Blanpain, C. Tracing the cellular origin of cancer. *Nat Cell Biol* **15,** 126–134 (2013).

5.Mamlouk, S. *et al.* DNA copy number changes define spatial patterns of heterogeneity in colorectal cancer. *Nat Comms* **8,** 14093 (2017).

6.Cheung, K. J. & Ewald, A. J. A collective route to metastasis: Seeding by tumor

cell clusters. *Science* **352,** 167–169 (2016).

The TP53 mutation database: a paradigm for the analysis of cancer genes

Thierry Soussi [1,2]

[1]INSERM, U1138, Centre de Recherche des Cordeliers, Paris, France, [2]Department of Oncology-Pathology, Karolinska Institutet, Cancer Center Karolinska (CCK) R8:04, Stockholm SE-171 76, Sweden

Purpose
Among the 14 million new cases of cancer diagnosed in 2012, 7 to 8 million tumours harboured a somatic *TP53* variant making *TP53* the most frequently mutated gene in human cancer. Germline *TP53* variants can also be identified in families prone to cancer such as Li-Fraumeni syndrome (LFS) as well as patients with early-onset breast cancer.

Methodology and results
We maintain the oldest TP53 mutation database that now includes 80,000 TP53 variants from patients with various types of cancer. Although most of these variants are somatic, some of them (about 1,000) are germline and generally associated with LFS. By using several data-driven methods on multiple independent quality criteria, we have been able to perform accurate curation of the database by eliminating the majority of sequencing artefacts and demonstrating that most of them were due to sequencing errors originating from poor quality DNA.
One of the major problems with the various variants concerns their pathogenicity and whether or not they are true driver variants, passenger variants or very infrequent germline polymorphisms [1]. One of the main problems related to the description of somatic variants is the lack of a true control as the tumour DNA sequence is usually compared to a reference using dbSNP to remove frequent and well characterized "neutral variants."

TP53 has been sequenced in more than 100,000 tumours, suggesting that it is one of the regions of the human genome that has been most frequently scrutinized for mutations and it is likely that some reported "infrequent somatic variants" are indeed very infrequent germline polymorphisms (less than 1/5,000 or 1/10,000).

Using *in silico* and functional analysis, we have conducted an analysis in order to track these "hidden variants", as they could lead to false clinical diagnoses and we have identified several new uncommon non-synonymous coding germline TP53 variants, some of which have been falsely designated as “driver variants” both in tumours and in families prone to cancer.

A combination of careful curation and detailed analysis of a specific Locus Specific Database (LSDB), i.e TP53, has led to the development of specific and accurate tools to predict the pathogenicity of each variant.

Conclusion
Our data point to a future in which the development of individualized predictive tools for each gene based on a combination of highly specific criteria will be necessary to meet clinical needs.

[1] Leroy, B., Ballinger, M. L., Bond, G. L., Braithwaite, A., Concin, N., Donehower, L. A., El-Deiry, W. S., Fenaux, P., Gaidano, P., Langerød, A., Hellstrom-Lindberg, E., Iggo, R., Lehmann-Che, J., Mai, P. L., Malkin, D., Moll, U. M., Myers, J. N., Nichols, K. E., Pospisilova, S., Ashton-Prolla, P., Rossi, D., Savage, S. A., Strong, L. C., Tonin, P. N., Zeillinger, R., Zenz, T., JF, F. J., Taschner, P. E. M., Hainaut, P., and Soussi, T. (2017). Moving TP53 mutation detection into clinical application: revised recommendations and guidelines. Cancer Res. *In press*,

THE HUMAN VARIOME PROJECT
an NGO official partner of UNESCO
sharing data • reducing disease

PRE-MEETING WORKSHOP

9.00 am - 10.40 am

5 June 2017

Variant nomenclature - How to describe sequence variants found?

The basis for reporting changes in a DNA sequence are the so called “HGVS recommendations for the description of sequence variants". In this workshop you get the chance to meet the expert and discuss the recommendations. After a short introduction and overview of the current recommendations the floor is yours. You can ask questions, discuss problems with the current recommendations, come with a practical example
and ask for advice, etc. Throughout the workshop, the expert will give multiple choice problems for you to solve and test your knowledge of the HGVS rules.

See: http://varnomen.hgvs.org/

PRE-MEETING SPONSORED WORKSHOP

9.00 am - 10.40 am

5 June 2017

Clinical Genomics - Interpretation and Reporting with QIAGEN Bioinformatics

We will be demonstrating how Ingenuity Variant Analysis (IVA) and QIAGEN Clinical Insight Interpret (QCI I) can help you overcome the challenges of prioritising, classifying and reporting variants. In the workshop we will be analysing a variety of different sample types and discussing how to:

- Generate a list of clinical relevant variants from WES and targeted panels
- Use the ACMG Guidelines for classification in an automated setting
- Associate variants with up-to date treatment and clinical trial information
- Generate a comprehensive clinical report

Please join us for this interactive hands-on session.

If you wish to try IVA during the workshop please bring a laptop: datasets and logon will details be provided.

Delegates are welcome to bring their own vcf files for analysis.

interactive biosoftware

SPONSORED WORKSHOP

2.10 pm - 3.40 pm

6 June 2017

The Alamut® Software Suite

The Alamut® Software Suite is a consistent set of applications dedicated to annotate, filter, and interpret genomic variations. Alamut Batch is a high-throughput annotation software for NGS analysis, Alamut Focus is an interactive variant filtration application, and Alamut Visual is a sophisticated gene browser designed to help assess the pathogenic status of genome variations in clinical and research settings.

In the Alamut workshop we will first introduce each software component, and then practice using Alamut Visual with several variants, including splicing mutations.
Attendees will be invited to download and install the software beforehand and to bring their laptop at the workshop.

Requirements:

- Attendees must bring their own laptop.
- Download instructions will be available one week before the meeting, please come prepared to make the most of this workshop.

THE HUMAN VARIOME PROJECT
an NGO official partner of UNESCO
sharing data • reducing disease

WORKSHOP

4.00 pm - 6.00 pm

6 June 2017

Genetic Variant Interpretation

Organizers: A Laner, R Hart, MJ Sobrido

While performance and output speed of sequencing platforms keep growing, so does the challenge of interpreting the significance of the list of genetic variants in a given individual. Integrative bioinformatic pipelines, accessible databases and *in silico* prediction tools facilitate this task. However, deciding about the potential pathogenicity of every variant is still an arduous work that requires a high level of expertise. In this workshop, examples of patients studied by NGS will be presented so that the attendees will have the opportunity to apply genetic variant filtering and interpretation protocols. The examples will include familial and sporadic cases of neurodegenerative disorders, intellectual disability and cancer. Proposed criteria, scoring systems and guidelines for variant interpretation will be discussed in order to address the main caveats of variant filtering and pathogenicity assessment.

Directory of Posters

M = Poster Session 1: Monday 5th June T = Poster Session 2: Tuesday 6th June		Presenting Author
POS001-M	Identifying placental eQTLs and determining biological importance of selected SNPs.	Triin Kikas
POS002-T	Multiplex SNP panels based on MALDI-TOF mass-spectrometry genotyping for association analysis of cognitive performance and Alzheimer's disease	Vadim Stepanov
POS003-M	Identification of genes harbouring rare variants of moderate to large effect in schizophrenia: a replication study	Julio Rodriguez
POS004-T	Cito-Nuclear Incompatibility As A Risk Factor For Neurodevelopmental Disorders	Javier Gonzalez-Peñas
POS005-M	Sequence level genotyping at the TCF4 CTG repeat associated with Fuchs endothelial corneal dystrophy	Mariam Alkhateeb
POS006-T	Somatic mutations are abundant in focal cortical dysplasia	Vanessa de Almeida
POS007-M	Association between genetic polymorphisms in RANK/RANKL/OPG signaling pathway with developmental enamel defects and dental caries experience	Erika Küchler
POS008-T	Association between genetic polymorphisms in DEFB1 and in miRNA202 with caries experience in children	Daniela Oliveira
POS009-M	Modeling atorvastatin plasmatic concentrations in healthy volunteers using integrated pharmacogenetics sequencing	Omar Cruz Correa
POS010-T	Mutational screening of desmosomal and non-desmosomal genes in Russian patients with arrhythmogenic right ventricular cardiomyopathy	Anna Shestak
POS011-M	Association with Idiopathic Pulmonary Fibrosis is attributable to common rather than rare variants in 11p15, 14q21, and 17q21	Jose Miguel Lorenzo Salazar
POS012-T	Identification of a novel CASQ2 deletion causing Catecholaminergic polymorphic ventricular tachycardia	Alejandro Blanco-Verea
POS013-M	Evaluation of genotype-phenotype relationship of short NOP56 expansions causing SCA36	María García-Murias
POS014-T	Triple Negatives in Myeloproliferative Neoplasms: is triple negative positive?	Annabel Kearney
POS015-M	Development of cost/effective strategies for genetic diagnosis of Polycystic Kidney Disease (PKD) based on the population mutagenesis rate or specific needs.	María Lara Besada Cerecedo
POS016-T	Mapping mitochondrial heteroplasmy in a Leydig tumor by laser capture micro-dissection and cycling temperature capillary electrophoresis	Paulo Refinetti

POS017-M	Identification of disease causing variants of Leukocytoclastic Vasculitis in related individuals	Clara Mulhern
POS018-T	Gene Panel testing Improves Genetic Diagnosis in Hereditary Breast / Ovarian Cancer	Ana Blanco
POS019-M	Germline Promoter Hypermethylation Analysis Of Brca1 And Brca2 Genes In Hereditary Breast And Ovarian Cancer Patients	Marta Rodriguez Balada
POS020-T	The Challenge Of Interpreting Heterozygous Carriers In Niemann-Pick Type C Disease	Cristina Castro Fernández
POS021-M	A Customisable Scripting System for Identification and Filtration of Clinically Relevant Genetic Variants in Whole Exome or Large Gene Panel Data	Robert Smith
POS022-T	Save variant annotation sharing across laboratories	Beat Wolf
POS023-M	HGVS Variant Descriptions for Short Tandem Repeat Structures	Johan T. den Dunnen for Jonathan Vis
POS024-T	Investigating ACMG rules and quantitative methods for TP53 variant classification	Cristina Fortuno
POS025-M	Variant detection, interpretation and reporting using an end-to-end analysis pipeline	Ruth Burton
POS026-T	Creating new national variation databases and their regulatory environments in Japan	Netsuke Yamamoto
POS027-M	A New National Initiative and Data Sharing Approach for Genomic Medicine in Japan	Jusaku Minari
POS028-T	Interpreting the functional effects of variants with integrated platforms from UniProt	Andrew Nightingale
POS029-M	In silico Prediction of Deleterious SNPs in Endocytosis Genes Implicated in Alzheimer's disease	ChongHan Ng
POS030-T	Application of fibre-FISH in the characterization of complex structural rearrangement and variation: an overview	Fengtang Yang
POS031-M	Overcoming the diagnostic challenges in neurological disorders: the role of next-generation sequencing	Sofia Gouveia
POS032-T	Next-generation sequencing meets splicing: a multiplexed minigene splicing assay for exonic variants of OPN1LW and OPN1MW	Elena Buena Atienza
POS033-M	Utility of trio exome sequencing as a first-line diagnostic test for neurodevelopmental disorders	Ana Fernández-Marmiesse
POS034-T	Massive parallel sequencing for universal indications: diagnostic variant detection in a regional genetic center	Elisabeth Maurer
POS035-M	Targeted Next- Generation Sequencing For Molecular Diagnosis Of Ichthyosis	Uxía-Saraiva Esperón-Moldes
POS036-T	Cycling Temperature Capillary Electrophoresis: A quantitative, fast and inexpensive method to detect mutations in mixed populations of human mitochondrial DNA	Paulo Refinetti

POS037-T	Germline mutations in childhood cancer patients suspected of genetic predisposition to cancer - a retrospective analysis	Dianne Sylvester
POS038-M	Dissecting the molecular basis of epileptic disorders in the Iberian Peninsula.	Sofia Gouveia
POS039-M	PGD counselling for variants of unknown significance	Ekaterina Pomerantseva
POS040-T	Comprehensive variant analyses including whole genome sequencing in hereditary colorectal cancer syndromes	Anna Rohlin
POS041-M	Evaluation Of Basic Massive Parallel Sequencing Parameters In Relation To True/False Positivity's Findings Of Rare Variants From Isolated Population With High Incidence Of Parkinsonism	Radek Vodicka
POS042-T	Potentially pathogenic germline CHEK2 and NOTCH3 variants among multiple early-onset cancer families	Men Dominguez Valentin
POS043-M	Allelic drop-out is a common phenomenon in the PCR-based NGS and emphasizes the importance of cross-validation.	Anna Bukaeva
POS044-T	Challenges of interpreting sequence variants: experience of a molecular laboratory with a panel of epileptic encephalopathy	Susana Sousa
POS045-M	Copy number variation differences between responders and non-responder to anti-TNF drugs in moderate-to-severe psoriasis	María C Ovejero-Benito
POS046-T	Identifying limiting factors for single nucleotide and copy number variant calling in a targeted sequencing panel	Pilar Cacheiro
POS047-M	PattRec: An easy-to-use CNV detection tool optimized for targeted NGS panels	Iris Roca
POS048-T	Determining the burden of copy number variation in patients with epilepsy	Tania de Araujo

The **Most Comprehensive View** of Genomes, Transcriptomes, and Epigenomes

MEET THE SEQUEL® SYSTEM, the scalable platform for SMRT® Sequencing. PacBio delivers long reads, high consensus accuracy, uniform coverage and epigenetic characterization.

HUMAN BIOMEDICAL RESEARCH

PLANT AND ANIMAL SCIENCES

MICROBIOLOGY AND INFECTIOUS DISEASE

Learn more at pacb.com/sequel

Poster Abstracts

POS001 - M

Identifying placental eQTLs and determining biological importance of selected SNPs

Triin Kikas[1]; Kristiina Rull[1, 2, 3]; Siim Sõber[1], Maris Laan[1,4]

1. *Human Molecular Genetics Group, Institute of Molecular and Cell Biology, University of Tartu (UT)*
2. *Department of Obstetrics and Gynaecology, UT*
3. *Women's Clinic, Tartu University Hospital*
4. *Institute of Biomedicine and Translational Medicine, UT*

Purpose: Expression quantitative trait loci (eQTL) are SNPs that can determine or alter the expression of certain genes; often the effect can be tissue-specific [1]. Cis-eQTLs are located in the close proximity of the interacting gene, whereas trans-eQTLs can be located anywhere else on the genome. So far placental eQTLs have been an area of minimal research. As of yet, no whole genome study of eQTL identification in placental tissue has been published. However, placental gene expression regulation, including eQTLs, is essential to a normal development and growth of a foetus as it has been shown that in preeclampsia and other pregnancy complications there are several changes in gene expression of placenta [2, 3]. Aim of this study is to conduct a whole genome screen to identify placental cis- and trans-eQTLs and study their genetic role as genetic risk factors for adverse pregnancy outcomes.

Methodology: First, a whole genome analysis on placental transcriptomes (n=40) was conducted to identify cis- and trans-eQTLs using linear regression. This analysis found 125 SNPs that were associated with 23 genes. Next, a validation for cis-eQTLs and selected top trans-eQTLs was carried out using Sequenom genotyping (n=300) and Taqman gene expression assay (24/each gene with 8 individuals for each genotype when possible). This part of the study utilized placental samples collected in the framework of REPROMETA study. One of the validated regions is currently under testing in a separate cohort collected 2012-2015 during Happy Pregnancy study. DNA extraction from 331 placental tissues and following genotyping with Taqman assay is currently on-going.

Results: In the discovery stage we found 125 robust SNPs (FDR corrected $P \leq 2.69 \cdot 10^{-4}$) that are associated with the placental expression of 23 genes. Follow-up stage focusing on the top 10% of the identified placental eQTLs, validated 7/11 SNPs, associated with three genes potentially linked to pregnancy complications. Testing of the relevance of these genetic variants as risk factors for adverse pregnancy outcome is currently ongoing in the Happy Pregnancy cohort.

Conclusions: Despite a modest sample size, we identified 125 eQTLs in placental tissue, out of which 7 were also experimentally validated in independent sampleset.

References:

1. McKenzie, M., Henders, A.K., Caracella, A., Wray, N.R., and Powell, J.E. (2014). Overlap of expression quantitative trait loci (eQTL) in human brain and blood. BMC Med. Genomics 7, 31.
2. Sõber, S., Reiman, M., Kikas, T., Rull, K., Inno, R., Vaas, P., Teesalu, P., Marti, J.M.L., Mattila, P., and Laan, M. (2015). Extensive shift in placental transcriptome profile in preeclampsia and placental origin of adverse pregnancy outcomes. Sci. Rep. 5, 13336.
3. Sitras, V., Fenton, C., and Acharya, G. (2014). Gene expression profile in cardiovascular disease and preeclampsia: A meta-analysis of the transcriptome based on raw data from human studies deposited in

Gene Expression Omnibus. Placenta 36, 170–178.

POS002 - T

Multiplex SNP panels based on MALDI-TOF mass-spectrometry genotyping for association analysis of cognitive performance and Alzheimer's disease

Vadim Stepanov[1,2], Kseniya Vagaitseva[1,2], Anna Bocharova[1], Andrey Marusin[1], Oksana Makeeva[3]

[1]Institute of Medical Genetics, Tomsk National Medical Research Centre, Tomsk, Russian Federation
[2]Tomsk State University, Tomsk, Russian Federation
[3]Nebbiolo LLC, Tomsk, Russian Federation

Purpose
Despite the progress in sequencing techniques, targeted SNP genotyping is still widely used in human and medical genetics. An analysis of genetic variability measured by restricted number (usually, several dozens) of targeted SNPs is a common approach in such basic areas as replicative association studies on common diseases, human evolutionary and population genetics, as well as in practical areas like DNA diagnostics and forensic DNA identification. Among other platforms designed for multiplex genotyping of several dozens of SNPs, MALDI-TOF DNA mass-spectrometry is one of the most powerful and cost-effective. In this paper we describe a development of multiplex MALDI-TOF based SNP panel for association analysis of human cognitive performance and Alzheimer's disease (AD).

Methodology
At first stage of marker selection, a list of SNPs associated according recent genome-wide association studies (GWAS) with cognitive traits (cognitive tests performance scores, intelligence, mathematical abilities etc.) was formed. On the next step, 177 SNPs associated with various cognitive functions in more than one GWAS, or SNPs associated both with cognitive traits and AD were selected. 13 of them demonstrating multiple associations were considered as priority SNPs. Finally, based on maximizing multiplexing capacity, primer's compatibility and SNPs prioritization, a panel of 62 SNPs in 48 genes, including 12 out of 13 priority SNPs, was formed. Genotyping of the panel was performed in two multiplexes (32 and 30 markers) by multilocus PCR and mass analysis of DNA molecules using Sequenom MassARRAY 4 (Agena Biosiences) platform. 62 markers were genotyped in two samples: 1) 708 elderly people (average age - 71.8 years), healthy against neurodegenerative and psychiatric diseases with normal cognitive status according MoCA score test; 2) sample of 218 patients with Alzheimer's disease (average age - 72.15 years). Both samples represent the Russian population of the city of Tomsk.

Results
In the analysis of variability of cognitive functions in sample of elderly people with normal cognitive status, 17 markers were detected associated with MoCA performance test, including genetic markers of APOE, TOMM40, PVRL2 and APOC1 genes, located in the same locus on chromosome 19q13.32. Highly significant association with MoCA was also found for gene markers of SORL and CSMD1 genes. Analysis of the association of genetic markers with the AD in the case-control design revealed 11 markers whose frequencies differ significantly in AD patients and healthy controls. 5 of SNPs at APOE-TOMM40 cluster as well as markers of CSMD1 gene demonstrate association both with normal cognitive performance in elderly and with AD.

Conclusion
Cross-replication analysis demonstrates that the genetic component of the normal variability of the cognitive functions in the

elderly and the genetic component of AD with late manifestation partially overlap. Multiplex SNP genotyping based on MALDI-TOF platform proved to be a useful tool for rapid and cost-effective small-scale association studies.
This work was supported by the Russian Science Foundation (project # 16-14-00020)

POS003 - M

Identification of genes harbouring rare variants of moderate to large effect in schizophrenia: a replication study

Julio Rodríguez-López[1], Beatriz Sobrino[1], Jorge Amigo[1], Mario Páramo[1], Manuel Arrojo[1], Javier Costas[1]

[1]Instituto de Investigación Sanitaria (IDIS) de Santiago de Compostela, Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Spain

Genetic architecture of Schizophrenia is complex and involves different genetic factors. Several studies showed that rare, putative damaging variants play an essential role in its genetic aetiology. Based on this criteria, Purcell et al. [1] studied 2,546 candidate genes in schizophrenia previously reported by several studies, looking for enrichment in:

- Rare (MAF<0.001) and disruptive mutations, and
- Rare (MAF<0.001) and missense variants rated as “damaging” by all of five prediction algorithms employed (PolyPhen2, HumDiv and HumVar, LRT, MutationTaster, and SIFT).

They reported this database in the genebook web server, http://research.mssm.edu/statgen/sweden/, and using these information, we elaborated a list of the 36 genes that showed a significant increase of those variants in schizophrenia cases compared with healthy controls (OR=3.5, 95%CI : 2.82-4.32). With the intention to replicate these results in our collection of samples, we carried out a genetic replication study, using Amplisec Next Generation Techniques to sequence these genes in 327 schizophrenia samples and 318 healthy controls (80% power to detect an OR of 2.3 at alfa level of 0.05). After applied the strict criteria used by Purcell et al. to detect putative functional variants in the risk to schizophrenia, our results did not replicate the Purcell’s data. However, exploratory analysis showed a nearly significant increase of disruptive variants in cases vs controls.
These results show that identification of candidate genes that could be used in the diagnosis and personalized treatment of schizophrenia through the analysis of rare variants, even using the most robust results published so far, presents a lot of caveats. Indeed, one of the most important handicaps is the enhancement of the bioinformatics tools to help the discrimination of the most putative damaging variants. The improvement of this specific step of analysis is essential to lead to the identification of truthfully candidate genes in schizophrenia.

References

1 Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P et al (2014) A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506:185–190. doi: 10.1038/nature12975

POS004 - T

Cito-Nuclear Incompatibility As A Risk Factor For Neurodevelopmental Disorders

J González-Peñas [1], E Gonzalez-Vioque [1] , B Arias [1], A Latorre-Pellicer [3] A Carracedo [2, 3], M Parellada [1]

[1]*Instituto de investigación Sanitaria Hospital Gregorio Marañón, Madrid, Spain*
[2]*Instituto de Investigación Sanitaria (IDIS) de Santiago de Compostela, Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Santiago de Compostela, Spain*
[3]*Universidad de Santiago de Compostela*

Coevolution explains the coordinated changes that occur in a pair of organisms or biomolecules throughout their evolution, generally to maintain or improve the functional interaction between them [1]. Recently, this concept has been extended to a molecular level [2].

Although coevolution of two different systems occurs in the context of numerous biological processes, the case of mitochondrial system is of special importance for several reasons: First, genes found in two physically separate genomes, nuclear (nDNA) and mitochondrial (mtDNA), are coordinately expressed and they cooperate for a proper mitochondrial function. Second, mtDNA accumulates mutations at a higher rate than nDNA because of several factors as less effective correction, absence of histonic or the physical proximity to points reactive oxigen species (ROS) production. Thus, nDNA has to evolve faster than expected in order to adapt to mtDNA changes, that is, to different mitochondrial haplotypes.

Due to this fast coevolution, potential incompatibility (mismatches) between nuclear and mitochondrial genomes from different populations of the same species could lead to a significant impact on human health. In fact, mito-nuclear incompatibility has been extensively demonstrated using different models and experimental approaches [3,4].

In this regard, it is necessary to emphasize the work carried out by Mishmar et al. focused on OXPHOS system [5], which has led to the identification of proteins encoded in both genomes that interact physically, reaching the identification of individual amino acids involved in this interaction. In a subsequent work, they described variations in this interaction, defined as coevolution disruptions, that are associated with human pathology [6]. Moreover, it has been recently proposed that different interactions of nuclear and mitochondrial genes can yield to significant changes of the electron transport chain, causing a particular metabolic adaptation [7]
As a significant proportion of patients with neurodevelopmental disorders have associated mitochondrial dysfunction but no studies have succeed to find pathological mutations in mtDNA or nDNA that could explain this dysfunction, we hypothesize that different adjustments of the OXPHOS system could condition the pathophysiology of some developmental disorders.
We performed a comparative analysis of the genetic variation in OXPHOS nuclear genes and characterized mitochondrial haplotypes from the 1000 genomes database (supposedly healthy population). A detailed statistical analysis was carried out, based on these data, in order to search for nuclear genetic variants coexisting with specific haplotypes. We ranked nuclear OXPHOS genes according to their pattern of coevolution, and highlighted those that explained to a greater extent the mitochondrial haplotype diversity.

Using a sample of more than 400 Spanish Autism trios, we will analyze the transmission of candidate nuclear variants and mitochondrial haplotypes performing transmission disequilibrium test (TDT). Replication of previous findings will be done with of Google's trios sample (https://www.mss.ng/). Interrelation with other variables as heteroplasmy and total antioxidant status (TAS) will also be performed.

References

[1] Pazos, F. & Valencia, A. Protein co-evolution, co-adaptation and interactions. *EMBO J* **27**, 2648-2655, doi:10.1038/emboj.2008.189 (2008).

[2] de Juan, D., Pazos, F. & Valencia, A. Emerging methods in protein co-evolution. *Nat Rev Genet* **14**, 249-261, doi:10.1038/nrg3414 (2013).
[3] Levin, L., Blumberg, A., Barshad, G. & Mishmar, D. Mito-nuclear co-evolution: the positive and negative sides of functional ancient mutations. *Front Genet* **5**, 448, doi:10.3389/fgene.2014.00448 (2014).
[4] Roubertoux, P. L. *et al.* Mitochondrial DNA modifies cognition in interaction with the nuclear genome and age in mice. *Nat Genet* **35**, 65-69, doi:10.1038/ng1230 (2003).
[5] Gershoni, M. *et al.* Coevolution predicts direct interactions between mtDNA-encoded and nDNA-encoded subunits of oxidative phosphorylation complex i. *J Mol Biol* **404**, 158-171, doi:10.1016/j.jmb.2010.09.029 (2010).
[6] Gershoni, M. *et al.* Disrupting mitochondrial-nuclear coevolution affects OXPHOS complex I integrity and impacts human health. *Genome Biol Evol* **6**, 2665-2680, doi:10.1093/gbe/evu208 (2014).
[7] Latorre-Pellicer, A. et al. Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature* **535**, 561-565 (2016).

POS005 - M

Sequence level genotyping at the TCF4 CTG repeat associated with Fuchs endothelial corneal dystrophy

M. A. Alkhateeb[1], S. Cumming[1], M. Ciosi[1], A. Maxwell[1], A. Alshammari[1], The Scottish Myotonic Dystrophy Consortium[1], Generation Scotland[2], G. Hamilton[3], D. G. Monckton[1];

[1]Institute of Molecular, cell and system biology, University of Glasgow, Glasgow, United Kingdom, [2]Institute of Genetics & Molecular Medicine, Western General Hospital, University of Edinburgh, Edinburgh, United Kingdom, [3]Glasgow Polyomics, Research Institute of Molecular, Cell & Systems Biology, University of Glasgow, Glasgow, United Kingdom.

Introduction: Fuchs endothelial corneal dystrophy (FECD) is an eye disease affecting the corneal endothelium. FECD affects about 5% of the population over 40 years of age. Late onset FECD is associated with the expansion of a CTG repeat (CTG18.1) in the 3rd intron of the transcription factor 4 (TCF4) gene on chromosome 18. The traditional method of sizing CTG18.1 alleles is limited due to the assumption that the length of the fragment is equal to the repeat number. This might not always be correct due to the presence of another repeat $(CTC)_n(CTT)_n(CTC)_n$ downstream of the CTG repeat. The traditional method cannot measure somatic mosaicism of the expanded CTG repeat and detect the presence of any variants.
Materials and Methods: next generation sequencing was used to genotype germline alleles. Locus specific primers were modified with barcoded Illumina adapters and sequenced by MiSeq platform. Small pool PCR was carried out to quantify somatic instability.

Results: blood DNA samples from 584 individuals from the Scottish population (1168 alleles) were sequencing. This revealed 1116 alleles have unexpanded alleles (≤40 CTG repeats) and 52 alleles have expanded alleles (>40 CTG repeats). There were two variants present within the CTG repeats (GTG) and (CTC) in unexpanded alleles. No variant was observed in the expanded alleles. Sequencing revealed that the first $(CTC)_{5-8}$ repeat immediately downstream of (CTG) was polymorphic. In addition, (CTT) and the second (CTC) were monomorphic so far.
Conclusions: CTG18.1 locus is somatically unstable. The degree of instability is correlated with allele size and age.

POS006 - T

Somatic mutations are abundant in focal cortical dysplasia

V.S. de Almeida[1], S.H. Avansini[1], M. Borges[1], F.R. Torres[1], F. Rogerio[2], B.S. Carvalho[3], F. Cendes[4], I. Lopes-Cendes[1]

[1]Department of Medical Genetics, School of Medical Sciences, [2]Department of Anatomical Pathology, School of Medical Sciences, [3]Department of Statistics, Institute of Mathematics, Statistics and Scientific Computing, IMECC, [4] Department of Neurology, School of Medical Sciences; University of Campinas, UNICAMP, Campinas, SP, BRAZIL and the Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, SP, BRAZIL.

Introduction: Malformations of cortical development (MCD), including focal cortical dysplasia (FCD), can cause epilepsy and are often associated with the occurrence of refractory seizures [1]. FCD is characterized by alterations in the cortical cytoarchitecture also observed in other MCDs, such as in Tuberous Sclerosis (TS) and Hemimegaencephaly (HME) [2,3]. Recently, mosaic mutations were detected in TS, HME and in rare cases of FCD [4]; however, it is still unclear if only a single mutation event may lead to FCD or whether somatic mosaicisms is a more widespread genetic mechanism [4].

Materials and Methods: Deep whole exome sequencing was performed on genomic DNA extracted from brain tissue resected by surgery (BTRS) and blood samples of 12 patients with FCD. We performed capturing and enrichment with Nextera® Expanded Kit (Illumina®). Samples were sequenced following a 200bp paired-end protocol in a Hiseq2500 (Illumina®) to achieve at least 200x of average coverage. We aligned sequences using BWA-MEM and performed realignment around SNPs and indels, quality recalibration and variant calling using the Genome Analysis Toolkit (GATK). We evaluated mosaicisms using Mutect2. Variants were classified as mosaic mutations when less than 20% of reads are not aligned to the human genome reference and are present only in BTRS. Variants were filtered prioritizing frameshift, missense, nonsense and splicing site mutations. In addition, we also focused in variants not described previously in the normal population (reference individuals) or variants whose minor allele frequency (MAF) is $\leq$ 0.01. Effect of variants was evaluated using Variant Effect Predictor (VEP).

Results: We identified potentially deleterious mosaic mutations in BTRS in **all** 12 patients analyzed. Mosaic mutations were identified in 28 genes belonging to the mTOR pathway and in 40 genes belonging to the TAU pathway. More interestingly, **all** 12 patients have somatic mutations in BTRS in at least two genes.

Discussion/Conclusion: Somatic mutations were identified in genes with functional roles and expression in the central nervous system. Somatic mutations were abundant in FCD tissue, affecting several genes from the mTOR and Tau pathways, and were detected in all samples examined. Therefore, somatic mosaicism seems to be a much more abundant and widespread genetic phenomenon than previously described in FCD.

References: [1] Kuzniecky RI. Epilepsia. 35 Suppl 6:S44-56, 1994. [2] Fauser S, Huppertz HJ, Bast T, et al., Brain 129:1907-16, 2006. [3] Mühlebner A, Coras R, Kobow K, et al., Acta Neuropathol 123:259-72, 2012. [4] Becker AJ, Urbach H, Scheffler B, et al., Ann Neurol 52:29-37, 2002.

Supported by: CEPID-FAPESP

POS007 - M

Association between genetic polymorphisms in RANK/RANKL/OPG signaling pathway with developmental enamel defects and dental caries experience

[1]*Erika Calvano Küchler*, [1]Léa Assed Bezerra da Silva, [1]Raquel Assed Bezerra da Silva, [2]Lívia Azeredo Antunes, [1]Carolina Maschietto Puncinelli, [1]Maria Cecília Gorita dos Santos, [3]Patrícia Hernandéz-Gatón,[2]Leonardo Santos Antunes, [1]Paulo Nelson-Filho,

[1]Department of Pediatric Dentistry – School of Dentistry of Ribeirão Preto, University of São Paulo, Brazil. [2]Department of Specific Formation, School of Dentistry, Fluminense Federal University, Brazil. [3]The Catalan Institute of Health (ICS), Barcelona University, Spain.

Purpose: Studies with animal models [1] and ex vivo studies [2] demonstrated that RANK, RANKL and OPG are involved in the enamel formation and mineralization. Therefore, the present study aimed to evaluate the association between genetic polymorphisms in RANK, RANKL and OPG with developmental defects of enamel (DDE) and caries experience.

Methodology: Two-hundred and eight children (3 to 13 years old) who received dental treatment at the Ribeirão Preto Dental School, University of São Paulo-Brazil were examined. Caries experience was evaluated according to the dmft and DMFT index and enamel defects were evaluated according to DDE index. Saliva samples were collected from each patient for DNA extraction. Participants were divided as 'caries free' and 'caries experience', and as 'with DDE' and 'without DDE' groups. Polymorphisms (rs3826620, rs9594738 e rs2073618) in RANK, RANKL and OPG were evaluated by real-time PCR (TaqMan™). The means of caries experience according to the genotype were performed using ANOVA and post-hoc Tukey. Chi-square or Fisher´s exact tests were used to compare allele and genotype distributions between patients groups. A significant level of 5% was used.

Results: The studied polymorphisms were in Hardy-Weinberg equilibrium. DDE was not associated with the polymorphisms rs3826620, rs9594738 and rs2073618 in RANK, RANKL and OPG ($p>0.05$). A borderline association was found between caries experience in permanent dentition and the polymorphism rs3826620 in RANK ($p=0.06$).
Conclusion: These results demonstrated that DDE was not associated with the polymorphisms rs3826620, rs9594738 e rs2073618 in RANK, RANKL and OPG and suggested that the polymorphisms rs3826620 in RANK could be involved in caries experience in permanent teeth.

This study was supported by The São Paulo Research Foundation (FAPESP) 2015/06866-5 and 2016/08149-1 (ECK)

References
1- Ohazama A, Courtney JM, Sharpe PT (2004). Opg, Rank, and Rankl in tooth development: co-ordination of odontogenesis and osteogenesis. J Dent Res. 83:241-4.
2- Sheng ZF, Ye W, Wang J, Li CH, Liu JH, Liang QC, Li S, Xu K, Liao EY (2010). OPG knockout mouse teeth display reduced alveolar bone mass and hypermineralization in enamel and dentin. Arch Oral Biol.55:288-93.

POS008 - T

Association between genetic polymorphisms in DEFB1 and in miRNA202 with caries experience in children

[1]*Daniela Silva Barroso de Oliveira*, [2]Léa Assed Bezerra da Silva, [2]Raquel Assed Bezerra da Silva,

[3]Andrea Lips, [4]Katharina Holanda Morant de Oliveira, [3]Lívia Azeredo Antunes, [3]Leonardo Santos Antunes, [2]Paulo Nelson-Filho, [5]Gutemberg Gomes Alves and [2]Erika Calvano Küchler,

[1] Department of Clinic and Surgery – School of Dentistry of Federal University of Alfenas, Brazil. [2]Department of Pediatric Dentistry – School of Dentistry of Ribeirão Preto, University of São Paulo, Brazil. [3]Department of Specific Formation, School of Dentistry, Fluminense Federal University, Brazil. [4]Department of Dentistry, Federal University of Sergipe, Brazil.[5]Clinical Research Unit, Fluminense Federal University, Niterói, Rio de Janeiro, Brazil.

Purpose: To evaluate the association between genetic polymorphisms in DEFB1 and in microRNA202 with caries experience in children.

Methodology: A total of 222 children (3 to 13 years old) from the Southeast of Brazil were included in the study. Caries experience was evaluated according to the dmft and DMFT index. Saliva samples were collected from each patient for DNA extraction. Participants were divided as 'caries free' and 'caries experience'. The polymorphisms (rs11362 and rs1799946) in DEFB1 and the polymorphism (rs12355840) in microRNA202 were evaluated by real-time PCR (TaqMan™). Chi-square or Fisher´s exact tests were used to compare allele and genotype distributions in the free, recessive and dominant model. A significant level of 5% was used.
Results: The studied polymorphisms were in Hardy-Weinberg equilibrium. The genotype and allele distribution of the polymorphism rs11362 in DEFB1 was associated with caries experience (p=0,043).

Conclusion: The polymorphism rs11362 in DEFB1 is associated with caries experience in children from the Southeast of Brazil.

This study was supported by The São Paulo Research Foundation (FAPESP) 2015/06866-5 (ECK)

POS009 - M

Modeling atorvastatin plasmatic concentrations in healthy volunteers using integrated pharmacogenetics sequencing

Omar Fernando Cruz-Correa[1], Rafael Baltazar Reyes León-Cachón[2,3], Hugo Alberto Barrera-Saldaña[2,4] & Xavier Soberón[1,5]

[1]Instituto Nacional de Medicina Genómica, Periférico Sur No. 4809, Col. Arenal Tepepan, Delegación Tlalpan, México, D.F. C.P. 14610, Mexico
[2]Facultad de Medicina, Universidad Autónoma de Nuevo León, Ave. Madero, Col. Mitras Centro, Monterrey, Nuevo León, C.P. 64640, Mexico
[3]Centro de Diagnóstico Molecular y Medicina Personalizada, Universidad de Monterrey, Ave. Ignacio Morones Prieto Pte. 4500, Col. Jesús M. Garza, San Pedro Garza García, Nuevo León, C.P. 66238, Mexico
[4]Vitagénesis, SA de CV., Col. Colinas de San Jerónimo. Monterrey, Nuevo León, C.P. 64630, Mexico
[5]Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Cuernavaca, Morelos, C.P. 62210, Mexico

Purpose: To use variants found by Next-Generation Sequencing to model atorvastatin plasmatic concentration profiles (Area Under the plasmatic concentration Curve, AUC) in healthy volunteers. Subjects & methods: A total of 60 healthy Mexican volunteers participating in a pharmacokinetics study of atorvastatin were enrolled in this study. Target enrichment libraries for Next-Generation Sequencing were prepared using a custom HaloPlex™ probe panel (Agilent Technologies,

CA, USA). Possible functional effects of the variants found by next-generation sequencing were evaluated through a combination of informatics tools and reported associations in the Pharmacogenomics Knowledgebase (PharmGKB). We used variants with a predicted or reported functional effect across 20 genes involved in atorvastatin metabolism to construct a regression model using a support vector approach with a radial basis function kernel to model AUC. We then refined this model in order to explain a greater extent of the variance by selecting a subset of variants according to their impact on the model.Results: The final support vector regression model using 60 variants (including one novel variant) explained 93.53% of the variance in atorvastatin AUC. Conclusion: In order to model atorvastatin's AUC, an integrated analysis of several genes known to intervene in the different steps of metabolism and response is needed and no single genetic variant determines the pharmacokinetic profile of atorvastatin on its own. However, further studies are required in order to assess the usefulness and clinical relevance of this model in independent population samples, especially in those with a significantly different genetic background.

POS010 - T

Mutational screening of desmosomal and non-desmosomal genes in Russian patients with arrhythmogenic right ventricular cardiomyopathy

Anna Shestak [1], Olga Blagova [2], Yulia Lutokhina [2], Yulia Frolova [1], Sergey Dzemeshkevich [1], Elena Zaklyazminskaya [1,3]

[1] *Petrovsky Russian Research Center of Surgery*
[2] *The Department of Cardiology of the V.N.Vinogradov Faculty therapeutic clinic, I.M. Sechenov First Moscow State Medical University*
[3] *Pirogov Russian National Research Medical University*
anna.shestak87@gmail.com

Purpose: Arrhythmogenic right ventricular cardiomyopathy (ARVC) is an inherited disease characterized by progressive fibro-fatty replacement predominantly in the right ventricle, leading to the arrhythmias and heart failure. Mutations in the genes encoding desmosomal and non-desmosomal proteins account about 65% of ARVC cases. Identification of mutation in the causative genes is considered as a major diagnostic criterion. The purpose of this study is to assess diagnostic yield of the genetic screening of the 16 genes and mutational spectrum in the Russian ARVC with and without LV dilation cohort.

Methodology: Clinical and instrumental examination and genetic counseling were performed for 45 ARVC probands (mean age 38.5 y.o.) with definite, borderline or possible diagnosis of ARVC [1]. Mutation screening of the *PKP2, DSG2, DSP, DSC2, JUP, TMEM43, TGFB3, PLN, LMNA, DES, EMD, SCN5A, LDB3, CTTNA3, CRYAB,* and *FLNC* genes was performed by high-throughput sequencing on PGM IonTorrent (with authomatically designed Ampliseq oligoprimers) combined with PCR-based Sanger sequencing. The gene panel covers 94 kb of coding and adjacent intronic area with 521 primer pairs. The pathogenecity of the new genetic variants was evaluated *in silico* by PolyPhen2, SIFT or NetGen2, BDGP tools. The cascade familial screening for family members of genotype-positive probands was provided upon request.

Results: The average coverage of the selected genes was 91,9%; average size of the amplicons required Sanger re-sequencing was 8,6 kb per patient.
We have identified 19 likely pathogenic variants in 20 index cases: 7 in the *PKP2* gene, 5 in the *DSG2* gene, 3 in the *DSP* gene, 1 in the *JUP* gene, 1 in the *TMEM43* gene (in

two probands), 1 in the *LDB3* gene, and 1 in the *SCN5A* gene. Truncated and splicing mutations account 35%. Mutation:VUCS ratio was 4:1. We have analyzed the mutation detection rate in the groups with definite, borderline and possible diagnoses of ARVC and the total contribution of genetic data in the disease diagnosis. Patients carrying these genetic variants had clinical manifestation at the third/fourth decade of life with high-grade ventricular arrhythmia; and 3 ICDs were implanted.

Conclusion: The average diagnostic yield of the genetic screening was 44% with our gene panel. The mutation detection rate in the groups with definite, borderline and possible diagnoses of the ARVC was 50%, 18% and 25% respectively.

1 - Marcus FI., McKenna WJ., Sherrill D., Basso C., Bauce B., Bluemke DA., Calkins H., Corrado D., Cox MG., Daubert JP., Fontaine G., Gear K., Hauer R., Nava A., Picard MH., Protonotarios N., Saffitz JE., Sanborn DM., Steinberg JS., Tandri H., Thiene G., Towbin JA., Tsatsopoulou A., Wichter T., Zareba W. (2010) Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J. 31(7):806-14. doi: 10.1161/CIRCULATIONAHA.108.840827.

Study was supported by Russian National Foundation grant №16-15-10421

POS011 - M

Association with Idiopathic Pulmonary Fibrosis is attributable to common rather than rare variants in 11p15, 14q21, and 17q21

José M. Lorenzo-Salazar[1], Shwu-Fan Ma[2], Justin M. Oldham[3], Imre Noth[2], Carlos Flores[1,4,5]

[1]Instituto Tecnológico y de Energías Renovables (ITER), Genomics Division, Santa Cruz de Tenerife, Spain; [2]Pulmonary & Critical Care Medicine, University of Chicago, Chicago, IL, USA; [3]Interstitial Lung Disease Program, University of California at Davis, Davis, CA, USA; [4]Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna, Santa Cruz de Tenerife, Spain; [5]CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain

Purpose: Idiopathic pulmonary fibrosis (IPF) is a low incidence, devastating disease with unknown aetiology and high mortality. A genome-wide association study (GWAS) completed by our group [1] have revealed common variants in three loci (11p15.5, 14q21.3, and 17q21.31) that associate with IPF susceptibility. Here we performed a fine-mapping study of these loci in order to provide further delineation of the likely causal variants.

Methodology: A total of 1.7 Mb region was sequenced in samples from 181 European American IPF patients using the Agilent SureSelect XT2 kit and 100 bases paired-end reads generated on HiSeq 2000. Association testing with IPF susceptibility was performed on post-quality-filtered variants detected by GATK using reference data from 501 unrelated Europeans as controls.

Results: Approximately 10^7 reads per sample were generated with an average of 100X depth coverage across the regions of interest. Association tests on the 10,245 biallelic single nucleotide polymorphisms with call rates >95% identified 36 variants reaching genome-wide significance ($p<5x10^{-8}$). The strongest significance was obtained for a promoter variant of *MUC5B* (rs35705950:G>T, $p=2.7x10^{-22}$). Haplotype analyses revealed no further evident effect of other variants in this region. Besides, we identified four novel variants associated in 14q21.3 and 17q21.31, but only one of them (NC_000014.8:g.

47751911C>A) was rare among controls (0.3%).

Conclusion: Our results provide limited support for the existence of rare variants with large effects in IPF susceptibility residing in these loci.

Funding: Pulmonary Fibrosis Foundation (Chicago, IL); Coalition for Pulmonary Fibrosis (San Jose, CA); Core Subsidy Mini Awards of the Institute of Translational Medicine and Clinical and Translational Science Award (UL1 RR024999); and RO1HL130796.

References:

1. Noth I, Zhang Y, Ma SF, et al (2013) Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. Lancet Respir Med 1:309-317

POS012 - T

Identification of a novel CASQ2 deletion causing Catecholaminergic polymorphic ventricular tachycardia

Blanco-Verea A[1,2,3], Ramos-Luis E[1,2,3], Álvarez-Barredo M[2], López-Abel B[1,3], García-Seara J[2], Rodríguez-Mañero M[2], Martínez-Sande L[2], Fernández-López A[2], Carracedo A[3], González-Juanatey JR[2], Brion M[1,2,3]

1. Xenética de Enfermidades Cardiovasculares e Oftalmolóxicas, Instituto de Investigación Sanitaria de Santiago de Compostela, Complexo Hospitalario Universitario de Santiago de Compostela, Santiago de Compostela, Spain.
2. Servicio de Cardiología, Hospital Clínico Universitario de Santiago de Compostela-SERGAS, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Santiago de Compostela, Spain.
3. Grupo de Medicina Xenómica, Instituto de Investigación Sanitaria de Santiago de Compostela-Universidade de Santiago de Compostela-Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain.
4. Unidad de Cardiología Pediátrica, Servicio de Pediatría, Hospital Clínico Universitario de Santiago de Compostela-SERGAS, Santiago de Compostela, Spain.

Purpose: Catecholaminergic polymorphic ventricular tachycardia (CPVT) is a genetically heterogeneous disease with two inherited forms identified: an autosomal dominant form associated with mutations in the gene encoding the cardiac ryanodine receptor (RYR2), and a recessive form associated with mutations in the gene encoding the cardiac isoform of calsequestrin (CASQ2). Mortality rate of CPVT reaches up to 50% by the age of 30 in untreated individuals; however the patients can be effectively treated by beta adrenergic blockers and/or implantable cardioverter defibrillator (ICD). Therefore, clinical genetic screening for early diagnosis is of great interest.

Methodology: After arriving at our cardiology service a patient with a history of syncopes from childhood and recovered from a cardiorespiratory arrest, he was diagnosed with a possible CPVT and an ICD was implanted due to his family history of sudden death in two brothers of 8 and 14 years. With the intention of confirming the diagnosis and finding genetic factors involved, the patient's DNA was sequenced for a panel of 28 genes related to primary electrical diseases. Detected variants were prioritized based on the likelihood of causality, according to the recommendations.

Results: Sequencing did not show any single nucleotide genetic variant with causality but the sequence review of the two main genes, RYR2 and CASQ2, allowed us to verify that

the patient did not present sequencing data of the last three exons of the CASQ2 gene. The deletion was confirmed with alternative methods based on conventional sequencing. The family study allowed demonstrating the implication of the deletion in the development of CPVT and sudden death, when presented in homozygosis.

Conclusion: For the first time, the development of CPVT and the risk of sudden cardiac death associated with the disease is explained by relatively large losses or gains of genetic material. Therefore, special care should be taken when applying massive sequencing technologies in genetic diagnosis, being also important to look for copy number variation.

POS013 - M

Evaluation of genotype-phenotype relationship of short NOP56 expansions causing SCA36

García-Murias M[1], Arias M[2], Yañez-Torregroza Z[3], Cacheiro P[1], García-Sancho C[4], García-Antelo MJ[4], Barros F[5], Pardo M[6], Sobrido MJ[1], Quintáns B[1].

[1] Grupo de Neurogenética, Instituto de Investigación Sanitaria (IDIS); Grupo de Medicina Xenómica; CIBERER-U711, Santiago de Compostela, Spain.
[2] Servicio de Neurología, Complejo Hospitalario Universitario de Santiago, Spain.
[3] Grupo de Neurogenética, Instituto de Investigación Sanitaria (IDIS), Santiago de Compostela (Spain); Universidad Simón Bolívar, Barranquilla, Colombia.
[4] Servicio de Neurología, Complejo Hospitalario Universitario de La Coruña, Spain.
[5] Fundación Pública Galega de Medicina Xenómica; CIBERER-U711, Santiago de Compostela, Spain.
[6] Servicio de Neurología, Hospital Comarcal de Valdeorras, El Barco de Valdeorras, Spain.

Purpose: Spinocerebellar ataxia 36 (SCA36) is an autosomal dominant ataxia with sensorineural hearing loss and motor neuron dysfunction, usually starting in the fifth or sixth decade. The cause is an intronic hexanucleotide expansion in *NOP56*, with 4-15 repeat normal alleles and over 650 repeats, full-penetrant alleles. The pathogenicity of short expansions is not well documented.

Methodology: 26 patients with full expansions and 3 patients with short expansions underwent neurologic examination and SARA scoring for the severity of ataxia. Genotyping was carried out by standard and repeat-primed PCR on peripheral blood DNA.
Results: A 64 yo man with disease onset at 55 had an abnormal 35 repeat allele and SARA score of 13. His 86 yo mother, also with a 35 repeat allele, reported disease onset around 50. Her current SARA score was 20. A maternal uncle, with an expanded allele of 49 repeats, had disease onset around 60, and SARA score of 14 at age 70. The age-adjusted SARA score of patients with short expansions was not significantly different from that of full-expansion patients. The family history did not disclose cases of non-penetrance.

Conclusion: Short NOP56 expansions (at least ≥35 repeats) are pathogenic. There is no significant difference in the clinical impact of short expansion versus fully expanded alleles. However, more cases with short expansions need to be studied to reach firm conclusions on genotype-phenotype relationships, and thus improve genetic counseling. Penetrance, somatic and gonadal stability of short expansions require further investigations.
Funding: PI12/00742.

POS014 - T

Triple Negatives in Myeloproliferative Neoplasms: is triple negative positive?

Anabel Kearney[1 2], Patricia Rebeiro, Joëlle Marivel & Silvia Ling

1. *Department of Haematology, SWS Pathology Service Liverpool NSW Australia*
2. *Ingham Institute of Advanced Medical Research, Liverpool, NSW Australia*

Purpose
Different forms of Philadelphia-negative myeloproliferative neoplasms (MPNs) have heterogeneous prognoses. Minimal clinical intervention controls the relatively benign polycythaemia vera (PV) while essential thrombocythaemia (ET) and primary myelofibrosis (PMF) require more challenging management. Individuals suspected of MPN are usually tested for diagnostic mutations; Janus kinase 2 exon 14 (JAK2 V617F), JAK2 exon12 and thrombopoietin receptor (MPL) for PV and the same markers plus calreticulin (CALR) for the latter two diseases.
Triple negative patients are those suffering from PV, ET or PMF but prove negative for mutations in the above three genes and are worthy of screening for potential biomarkers that may predict an overall lower survival rate and, in particular, increased risk of leukaemic transformation in PMF [1]. The aim of the project is to establish whether common mutations of high risk myeloid neoplasm (acute myeloid leukaemia and myelodysplasia) are predictive of poor prognosis in triple negative MPN.

Methodology
Over a period of years, we have tested a large MPN cohort of individuals in our molecular laboratory. Upon presentation for possible MPN diagnosis, all patients are screened sequentially for JAK2 V617F, CALR and MPL or JAK2 exon12, according to their clinical progression. During this period of time, a subgroup of 20 triple negative patients has been identified.
Biomarkers thought to be of particular relevance to the screening of this group of individuals, especially those who may continue to transform to leukaemia, are tet methylcytosine dioxygenase 2 (TET2), isocitrate dehydrogenase (NADP+) 1 /2 (IDH1/2), nucleophosmin (NPM1), fms related tyrosine kinase 3 internal tandem repeat (FLT3-ITD), Wilms tumour 1 (WT1) and CCAAT/enhancer binding protein alpha (CEBPA) [2, 3]. Mutations have been detected by various methods including Sanger sequencing, and capillary electrophoresis.

Results
Mutation status of all triple negative patients is correlated with clinical history, survival and complications.

Conclusion
Currently there are no specific biomarkers for triple negative MPN. Therefore it is important to evaluate whether known mutations in other high risk myeloid disease could predict outcome of this subgroup of MPNs.

[1] Tefferi A, *et al.* (2014) Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, polycythemia vera and myelofibrosis. Blood 124:2507-2513.
[2] Lundberg P, *et al.* (2014) Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. Blood 123:2220-2228.
[3] Tefferi A (2010) Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms:JAK2, MPL, TET2, ASXL1, CBL, IDH, IKZF1. Leukemia 24:1128-1138.

POS015 - M

Development of cost/effective strategies for genetic diagnosis of Polycystic Kidney Disease (PKD) based on the population mutagenesis rate or specific needs

Lara Besada Cerecedo[1], Beatriz Sobrino[2], Jorge Amigo Lechuga[2], Patricia Regueiro Casuso[1], Ana María Barcia de la Iglesia[1], Manuel Fidalgo[3], Carmen Vázquez[1], Ángel Carracedo[2], Cándido Díaz Rodríguez[1] y Miguel A. García-González[1]

1 Group of Genetics and Developmental Biology of Renal Diseases. Health Research Institute of Santiago de Compostela (IDIS), Department of Nephrology, University Hospital Complex of Santiago de Compostela, Spain. 2 Galician Public Foundation of Genomic Medicine, Santiago de Compostela, Spain.

Purpose

Genetic tests have the benefit of ensuring an accurate diagnosis and anticipate the disease, yet the limitation of a high cost to be used in diagnostic routine. With the incorporation of the new generation sequencing (NGS), genetic tests have reduced dramatically the costs getting closer to the diagnosis by magnetic resonance imagen (MRI), with the advantage that the genetic diagnosis is one time for life for a ADPKD patient and 100 times cheaper for the genetic diagnosis of the other members of the family, in a disease where the type of mutation is critical for the prediction of disease progression and possible treatment decision making.

Methodology

Our group has recently solved one of the limitations of the application of NGS in the analysis of diseases with high homology pseudogenes (such as the *PKD1* gene). We developed an efficient genetic test for all cystic diseases: TEST-1) panel for the genetic region responsible of the most polycystic kidney disease in the population (the replicated region of *PKD1,* exons 1-34), TEST-2) Panel for common polycystic disease (with the 8 mutated genes most prevalent), and TEST-3) Panel for common, rare and ultra-rare polycystic kidney disease (all 72 cystic genes).

Results

By analyzing a total of 252 families clinically diagnosed for PKD, we end developing a strategy that identified the associated mutation in 90 families by using the TEST-1, 128 using TEST-2 and 34 using TEST-3. In addition we reanalyzed according the pathogenic score all genetic variants identified to the moment, that help us to establish a PKD database of a total number of 3260 reclassify variants in four categories: 1174 class-I (832 *PKD1*, 155 *PKD2* and 187 *PKHD1)*, definitely pathogenic (“stop codon”, “frameshift insertions”,“frameshift deletions” and “canonical splice site alterations”); 141 class-II (107 *PKD1*, 12 *PKD2* and 22 *PKHD1)*, probably pathogenic (“inframe deletion/ insertions”,“non canonical splice site mutations” and “amino acid substitutions"); 1594 class -III (1119 *PKD1*, 85 *PKD2* and 390 *PKHD1)*, of uncertain significance and 351 SNPs (199 *PKD1*, 21 *PKD2* and 131 *PKHD1,)* not associated with pathogenicity. A total of 29 *PKD1* variants, 8 *PKD2* and 15 *PKHD1* were novel PKD mutations. We extended the genetic analysis to all family members (n=2150 patients, affected and unaffected), what represents the expected 83% of the PKD Galician Population (Spain). Since, January 2016, the Galician Society of Nephrology will promote the genetic and phenotypical characterization of all expected PKD patients before the end of 2019, in a cost effective manner.

Conclusions

Here we describe the first cost/effective strategy applied for the diagnosis of all patients belonging to Local Health System for ADPKD.

POS016 - T

Mapping mitochondrial heteroplasmy in a Leydig tumor by laser capture micro-dissection and cycling temperature capillary electrophoresis

Paulo Refinetti[1], Stephan Morgenthaler[1] and Per Olaf Ekstrøm [2]

Chair of Applied Statistics [(1)], EPFL – FSB – STAP, 1015 Lausanne, Switzerland. Department of Tumor Biology [(2)], Institute for Cancer Research, The Norwegian Radium Hospital, Oslo, Norway.

Background: The growth of tumor cells is accompanied by mutations in nuclear and mitochondrial genomes creating marked genetic heterogeneity. Tumors also contain non-tumor cells of various origins. An observed somatic mitochondrial mutation would have occurred in a founding cell and spread through cell division. Micro-anatomical dissection of a tumor coupled with assays for mitochondrial point mutations permits new insights into this growth process. More generally, the ability to detect and trace, at a histological level, somatic mitochondrial mutations in human tissues and tumors, makes these mutations into markers for lineage tracing.

Method: A tumor was first sampled by a large punch biopsy and scanned for any significant degree of heteroplasmy in a set of sequences containing known mutational hotspots of the mitochondrial genome. A heteroplasmic tumor was sliced at a 12µm thickness and placed on membranes. Laser capture micro-dissection was used to take 25000µm^2 subsamples or spots. After DNA amplification, cycling temperature capillary electrophoresis (CTCE) was used on the laser captured samples to quantify mitochondrial mutant fractions.

Results: Of six testicular tumors studied, one, a Leydig tumor, was discovered to carry a detectable degree of heteroplasmy for two separate point mutations: a C → T mutation at bp 64 and a T → C mutation found at bp 152. From this tumor, 381 spots were sampled with laser capture micro-dissection. The ordered distribution of spots exhibited a wide range of fractions of the mutant sequences from 0 to 100% mutant copies. The two mutations co-distributed in the growing tumor indicating they were present on the same genome copies in the founding cell.
Conclusion: Laser capture microdissection of sliced tumor samples coupled with CTCE-based point mutation assays provides an effective and practical means to obtain maps of mitochondrial mutational heteroplasmy within human tumors.

POS017 - M

Identification of disease causing variants of Leukocytoclastic Vasculitis in related individuals

Ciara Mulhern[1], Ying Hong[1], Paul Brogan[1], Despina Eleftheriou[1]

UCL Great Ormond Street Institute of Child Health (ICH), London, UK

Introduction: Leukocytoclastic vasculitis is an inflammatory disease of the blood vessel wall, caused by neutrophilic infiltration and resulting in fibrinoid necrosis [2]. Its clinical presentation can vary significantly depending on size of affected vessels, extent of vascular injury, and whether an underlying pathology is also present [5]. Aetiology for the condition can vary from idiopathic, infection, drug or stress related [4] and many different subtypes of vasculitis exist. Its diverse phenotype presents a diagnostic challenge for clinicians. However, when it occurs in related individuals this can be indicative of an underlying genetic contribution. Siblings of consanguineous families typically share large

amounts of their DNA. The coefficient of inbreeding can be as high as 12.5% for first cousin marriages [1]. Geneticists often use this principle of autozygosity when hunting down disease causing variants [3,6]. Here we demonstrate the use of homozygosity mapping and whole exome sequencing in the identification of possible genetic permutations leading to the disease in a consanguineous family from Somalia.

Objective: To use homozygosity mapping and whole exome sequencing to discover genetic variants causative of leukocytoclastic vasculitis in the presenting family.

Patients and Methods: Three children aged 5, 8 and 9 from a consanguineous family in Somalia presented with recurrent cutaneous leukocytoclastic vasculitis. All in addition have severe learning difficulties, skin laxity and hypermobility, hyperpigmentation and spontaneous bruising. Skin biopsy in the youngest child confirmed neutrophilic leukocytoclastic vasculitis. Homozygosity mapping was performed on all family members.

Results and Future Directions:
Homozygosity mapping revealed large areas of autozygosity present in the genomes of the affected children. Identical by descent (IBD) regions which are common in all affected siblings but are not present in healthy individuals were used to create a candidate gene list. Whole exome sequencing will also be performed. Genes identified in both techniques become areas of high priority, when analysing the data and identifying a variant, causative of familial vasculitis.

1. Alkuraya F.S.(2010) One more tool in the clinical geneticists toolbox. Genet Med 2010:12(4);236-239
2. Arora A., Wetter D, Gonzalez-Santiago T et al (2014) Incidence of Leukocytoclastic vasculitis 1996-2010: A population based study in Olmsted county, Minnesota. Mayo Clin Proc, 2014 November; 89(11): 1515-1524.doi:10.1016/j.mayocp.2014.04.015
3. Frosk et al 2017 A truncating mutation in CEP55 is the likely cause of March, a novel syndrome affecting neuronal mitosis. J Med Genet 2016 doi. 10.1016/jmedgenet-2016-104296
4. Johnson EF, Wetter D.A, Lehman J.S.,Hand J.L. et al (2016) Leukocytoclastic vasculitis in children, Clinical characteristics, subtypes, causes and direct immunofluorescence findings of 56 - biopsy confirmed cases. J Eur Academy of Dermatology and Venerology 2016
5. Weiss P. (2012) Pediatric Vasculitis. Pediatr Clin North Am 2012 April ; 59(2):407-423 doi 10.1016/j.pcl.2012.03.013
6. Wang et al 2016 Homozygosity mapping and Whole genome sequencing links a missense mutation in POMGNT1 to autosomal recessive retinitis pigmentosa. Inv Opthal Vis Sci July 2016, Vol.57, 3601-3609. doi: 10.1167/iovs.16-19463

POS018 - T

Gene Panel testing Improves Genetic Diagnosis in Hereditary Breast / Ovarian Cancer

Ana Blanco1, Marta Santamariña2, Sandra Filippini3, Belinda Rodriguez2, Pablo Runa1, Uxía Esperón2, Ana Vega1

1 Fundacion Publica Galega de Medicina Xenomica (FPGMX), Hospital Clinico Santiago de Compostela, c/Choupana s/n, Santiago de Compostela, Spain
2 Genomic Medicine Group-USC, CIBER-ER, IDIS, Santiago de Compostela, Spain
3 Genetic Unit, Institute of Legal Medicine and Genomic Medicine Group, USC Faculty of Medicine, Spain

Purpose

Germline *BRCA1* and *BRCA2* pathogenic variants confer high breast/ovarian cancer risk. Since the identification of BRCA1 and BRCA2 genes in the ninties, PCR-based single gene analysis has been the predominant test for inherited breast and ovarian cancer (HBOC). However *BRCA1/ BRCA2* pathogenic variants are indentified only in 20-40 % of the HBOC families. Research efforts during this last decade lead to the identificationn of several new high and intermediate risk genes.
Recently, the applications of next- generation sequencing (NGS) technology have led to to the introduction of multiplex gene- panel testing for hereditary cancer.
The general objective of our work is to evaluate the sensitivity of the implementation of a NGS panel with 14 genes in the diagnosis of HBOC.

Methodology

In our laboratory we have developed a targeting resequencing multigenic panel of 14 genes involved in HBOC cancer and have incorporated it into the diagnostic routine. Included genes are: *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *MLH1*, *MSH2*, *MSH6*, *PALB2*, *PTEN*, *RAD51C*, *RAD51D*, *STK11*, *TP53*.

Results

To date we have analyzed more than 500 families and identified pathogenic variants in different genes, including: *CHEK2*, *PALB2*, *RAD51C*, *RAD51D*, *STK11* and *TP53*.

Conclusion

The implementation of multigene- panel using NGS technology increse the sensitivity of the diagnostic test for hereditary breast / ovarian cancer.

POS019 - M

Germline Promoter Hypermethylation Analysis Of BRCA1 And BRCA2 Genes In Hereditary Breast And Ovarian Cancer Patients

Rodríguez-Balada M[1], Roig B[1], Melé M[1], Salvat M[1], Martorell L[2], Atanesyan L[3], Borràs J[1], Gumà J[1].

[1] Cancer Genetic Counseling Unit (Oncology Research Group), Institut d'Oncologia de la Catalunya Sud (IOCS), Hospital Universitari Sant Joan de Reus, IISPV, Universitat Rovira i Virgili, Av. Del Dr. Josep Laporte, 43204, Reus (Spain).
[2] Hospital Universitari Institut Pere Mata, IISPV, Universitat Rovira i Virgili, CIBERSAM, C/Sant Llorenç, Reus, Spain.
[3] MRC-Holland b.v., Willem Schoutenstraat 1, 1057DL Amsterdam, the Netherlands.

PURPOSE: Germline mutations in the tumour suppressor genes *BRCA1* and *BRCA2* are associated with an increased risk of breast and ovarian cancer. Only 5-10% of hereditary breast and ovarian (HBOC) patients are carriers of *BRCA1* and *BRCA2* pathogenic mutations [1]. These heritable, non genetic changes, called epimutations could be germline, although only constitutional has been well established as a mechanism for cancer predisposition [2,3]. Aberrant methylation of *BRCA1* promoter regions is an alternative event of gene silencing in breast cancer. It has been described that inactivation of *BRCA1* by genetic or by epigenetic changes are mutually exclusive events [4, 5]. Also, Esteller et al, discarded this event as a second hit in BRCA mutated genes. With this in mind, the present study aimed to analyse the germline *BRCA1* and *BRCA2* promoter hypermethylation status of HBOC patient carriers and non-carriers of *BRCA1* and *BRCA2* pathogenic mutations and, if possible, to consider this epigenetic modification

analysis as an effective pre-screening tool for whole BRCA genetic analysis of HBOC patients [6].

METHODOLOGY: We analysed peripheral blood cell DNA of 96 HBOC patients, of which 14 were carriers of *BRCA1* pathogenic mutations, 9 were carriers of pathogenic mutations in *BRCA2* gene, 68 were non-carriers of pathogenic mutations in both genes and 5 were healthy control subjects. *BRCA1* and *BRCA2* Promoter Hypermethylation was assessed using a commercial SALSA MS-MLPA probemix ME0053-X1 according to manufacturer's protocol (MRC-Holland, The Netherlands) [7]. DNA fragments were analysed using an ABI 3500 Sequencer (Life Technologies, Spain). MLPA analysis was performed using the Coffalyser program. Healthy subjects results were used to normalize results data obtained from samples.

RESULTS: According to previous results, none of the HBOC patient carriers of pathogenic mutations in *BRCA1* or *BRCA2* genes showed promoter hypermethylation of *BRCA1* and *BRCA2* genes. In addition, none of the 68 HBOC patients non-carriers of *BRCA1* or *BRCA2* pathogenic mutations analysed showed promoter hypermethylation of *BRCA1* and *BRCA2* genes.

CONCLUSION: Our study showed no evidence of the possible contribution of *BRCA1* and *BRCA2* promoter hypermethylation analysis in the peripheral blood DNA of HBOC and therefore**,** it can not be considered as a useful pre-screening tool for BRCA genetic analysis of HBOC patients.

REFERENCES

[1] Boyle P, Ferlay J. (2005) Cancer incidence and mortality in Europe, 2004. Annals of Oncology 16(3): 481-488.
[2] Jones P, Baylin S. (2002) The fundamental role of epigenetics events in cancer. Nat Rev Genet 3 (6): 415-428.
[3] Horsthemke B. (2007) Heritable germline epimutations in human. Nat genet 39(5): 573-574.
[4] Esteller M, et al. (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. J. Natl Cancer Inst 92(7) : 564-9
[5] Bosviel R et al. (2011). BRCA1 promoter methylation in peripheral blood DNA was identified in sporadic breast cancer and controls. Clin Chim Acta. 412 (15-16) : 1472-5
[6] Esteller M et al. (2001). DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet. 10 (26): 3001-3007
[7] Nygren AOH et al. (2005). Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res 33(14): e128

POS020 - T

The challenge of interpreting heterozygous carriers in Niemann-Pick type C disease

Cristina Castro-Fernández[1,2,7], Manuel Arias[1,5], Víctor Rodríguez-Sureda[4,6], María García-Murias[1,2,6], Tania García-Sobrino[5] Carlos García-Sancho[3], Ángel Sesar[1,5], Celia Pérez-Sousa[3] , Beatriz Quintáns[1,2,6] Patricia Blanco-Arias[1,2] , Carmen Domínguez[4,6], María-Jesús Sobrido[1,2,6].

[1]Instituto de Investigación Sanitaria de Santiago, [2]Fundación Pública Galega de Medicina Xenómica, [3]Complexo Hospitalario Universitario de A Coruña, [4]CIBBIM – Nanomedicina , Vall d'Hebron Institut de Recerca (VHIR), [5]Complexo Hospitalario Universitario de Santiago, [6]CIBER de Enfermedades Raras, [7] Universidad de Santiago de Compostela.

Purpose: Niemann-Pick type C disease (NPC) is an autosomal recessive, lysosomal storage disorder caused by mutations in either *NPC1*

(NM_000271) or *NPC2* (NM_006432). Whether heterozygous variants can be pathogenic or confer susceptibility to overlapping neurodegenerative phenotypes is controversial. The aim of our study was the clinical, genetic and biochemical characterization of four patients with neuropsychiatric symptoms who carry heterozygous rare *NPC1* or *NPC2* variants.

Methods: Inclusion criteria for genetic screening of NPC were established by an expert group of neurologists. Screening of mutations in *NPC1* and *NPC2* was performed through a gene panel using the IonPGM platform and MLPA. We revised the detailed clinical records and comprehensive neurological examination of patients with potentially pathogenic variants. Filipin staining was carried out on cultured fibroblasts and biochemical markers were measured in peripheral blood.
The patients showed combination of intellectual disability, cognitive decline (frontal-subcortical profile), supranuclear gaze palsy, parkinsonism, dystonia and ataxia. No gene dosage abnormalities were identified through MLPA.

Conclusions: Although a pathogenic role of heterozygous variants has not been clearly established, it is plausible that some of these variants have an impact on the protein function and/ or a synergistic effect with variants in other genes, contributing to late-onset neurodegenerative manifestations of the FTD-parkinsonism spectrum. It is important to pursue further investigation of patients with rare heterozygous *NPC1/NPC2* variants and to study their effect on cellular pathways of lipid metabolism.

POS021 - M

A Customisable Scripting System for Identification and Filtration of Clinically Relevant Genetic Variants in Whole Exome or Large Gene Panel Data

Robert A. Smith*, Miles Benton*, Neven Maksemous*, Cassie Albury*, Deidre Roos-Araujo*, Larisa Haupt*, Rod Lea*, Lyn R. Griffiths*

** Genomics Research Centre, Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, Queensland, Australia.*

Purpose: Following recent developments in sequencing technology and the introduction of large scale next generation sequencing into research and clinical practice, new challenges have appeared in the detection of potentially pathogenic genetic variants and their use in diagnostics. This paper describes a method for the rapid filtration of variants from large next generation sequencing datasets to allow isolation of clinically relevant candidates for further verification in a diagnostic setting.

Methodology: Utilising a sequencing method agnostic system of public and open source databases, our method aligns and annotates sequencing data and queries variants using multiple *in silico* variant effect predictors. Variants are then sorted on the basis of simple, easy to modify gene lists for the target disease that give both flexibility of analysis and allow data masking to prevent incidental findings. Genes within the lists can be ranked by relationship to the desired disease or phenotype. The method outputs a standardised format written report of the most clinically relevant variants based on gene, variant type, variant population frequency and *in silico* predictions. Quality control metrics and analysis data specific to each run are recorded and saved. The method also

includes an integrated variant browser than can be used to quickly filter detected variants by multiple parameters and can be used to compare multiple samples, as well as view sequencing read information from the original data.
Results: This method as validated using 8 individuals previously sequenced for a neurological disorder, with a known pathogenic mutation. In all cases, the method returned the pathogenic variation in the top ranked variants for consideration, and all known variants in these individuals were returned by the method.

Conclusion: Our method represents a useful way of filtering large amounts of genome/exome data for diagnostic testing or screening, allowing human operators to quickly parse this data for clinically relevant variants for further consideration, especially in conditions where large numbers of genes could potentially affect observed patient phenotypes.

POS022 - T (*also an oral presentation Session 5*)

Safe variant annotation sharing across laboratories

Beat Wolf 1,2, Pierre Kuonen 1, David Atlan 3, Jonathan Stoppani 1, Davide Mazzoleni 1, Thomas Dandekar 2

1HES-SO Fribourg, Switzerland
2University of Würzburg, Germany
3Phenosystems SA, Belgium

Purpose
Annotating genetic variants heavily benefits from the usage of databases that already document the concerned variants. Due to the massive amount of possible variants and their complex effects, data sharing is essential to improve understanding of their consequences. Unfortunately many variants are only documented in private databases, either because of legal restrictions or because of the lack of time to publish the findings on public databases.
To solve this problem, we want to extract the information from where it already exists, which are the NGS data tools used in various laboratories. They often already integrate an internal variant database. We propose to augment those tools in a way to make it possible to share the variant annotations anonymously and securely among other users of the software, while giving the users full control on how to share their data.

Methodology
For our first prototype we choose to extend the graphical NGS data analysis software GensearchNGS, developed by Phenosystems SA, with the ability to share variant annotations across multiple laboratories. Our approach is based on an original concept we call Trusted-Friend-Computing (TFC). The basic idea of THC is the ability to share resources through a network of trusted friends. To implement this concept in GensearchNGS, we extend the open source POP-Java programming language with the required functionalities. POP-Java is a Java language extension that dramatically eases development of distributed Java applications. The TFC concept also allows us to explore the possibility to share computing resources across multiple laboratories, but, in the current project, we will focus on variant annotation sharing.

Results
We present the prototype of the proposed functionality in GensearchNGS as well as its open source implementation in POP-Java. We present the current state of the prototype as well as the underlying design principles.

Conclusion
Sharing variant annotations is essential to improve the quality NGS data analysis. We acknowledge the technical and legal restrictions which hinder wide spread sharing, and propose thus a more controlled way to share variant annotations among a group of

trusted laboratories. We hope that our work can serve as a step

POS023 - M

HGVS Variant Descriptions for Short Tandem Repeat Structures

Jonathan Vis[1] and Jeroen Laros[1]

Dept. of Human Genetics[1], Leiden University Medical Center, Leiden, Nederland
corresponding author: j.k.vis@lumc.nl

Introduction
A considerable part of the genome consists of repeated regions that occur in many forms. Here, we focus on *Short Tandem Repeats* (STRs). STRs contain *units* ranging from 2–5 base pairs which are repeated in *tandem*, i.e., immediately adjacent, typically 5–50 times. The combination of the lengths of different STRs is used in forensics for identification purposes. We present a method based on the recommendations of the Human Genome Variation Society (HGVS) [1] for the automatic construction of descriptions of repeated structures. As an additional requirement, the STR allele descriptions need to maintain compatibility with existing databases including the definition of the repeat units. In [2] three possible methods for describing STR alleles are envisioned: 1) complete sequence strings 2) a bracketed description 3) unique identifiers.

Method
The method consists of three parts: 1) finding repeat units 2) reference-based description of the repeat structure 3) relative description of the flanking regions.

Finding repeat units We adapted the run-length encoding algorithm for the detection of variable length repeated sequences in tandem. Care has to be taken when dealing with self-similar repeat units, e.g., TATA, because a smaller unit exists; TA. In these cases we are interested in the smallest unit.

Reference-based description of the repeat structure The purpose of this part of our method is to yield, given a set of repeat units, a bracketed type of description. The method expects a reference string as well as the observed string. The extraction of the description follows the method described in [3] with an alteration; we introduce a *masking* character, i.e., a character that does not match any character. After the masked variant extraction, substring matching is used to fill in the masked regions with repeat units from the repeat unit set.

Relative description of the flanking regions The final part deals with the automatic generation of relative descriptions for describing variation in the flanking regions. These relative descriptions are similar to coding region oriented descriptions used in HGVS for the description of 5' and 3' UTR variants, e.g., c.-56C>T and c.*32G>A. Mutalyzer [4] performs conversions between positioning schemes in HGVS. For our purpose we make use of this functionality by introducing an artificial coding region that represents the repeat structure.

Results
The descriptions resulting for our proposed method look like: TPOX(122_142):[-50G>T;TGAA(6)], where the selected ‘transcript' is made explicit in the notation of the reference sequence (TPOX). The repeat structure is selected from position 122 to 142. In the flanking region there is a single nucleotide substitution 50 bases upstream of the start of the repeat structure. The flanking region downstream of the repeat structure contains no variants.

References
[1] Sequence Variant Nomenclature, varnomen.hgvs.org, retrieved February 2017.
[2] K.B. Gettings, et al. “STR allele sequence variation: Current knowledge and future issues.” *Forensic Sci. Int.: Genetics* 18 (2015).
[3] J.K. Vis, et al. “An Efficient Algorithm for the Extraction of HGVS Variant Descriptions from Sequences.” *Bioinformatics* 31(2015).

[4] Mutalyzer 2.0.23, mutalyzer.nl, retrieved February 2017.

POS024 - T

Investigating ACMG rules and quantitative methods for TP53 variant classification

Cristina Fortuno[1], Paul James[2], Jue-Sheng Ong[1], Erin Young[3], Sean V. Tavtigian[3], Amanda B. Spurdle[1], ClinGen*TP53* Expert Panel Collaborators

[1]QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston QLD 4006
[2]Peter MacCallum Cancer Centre and Royal Melbourne Hospital Familial Cancer Centre
[3]Department of Oncological Sciences and Huntsman Cancer Institute, University of Utah School of Medicine

TP53 pathogenic germline variants cause Li-Fraumeni (LFS) and Li-Fraumeni-like (LFL) syndromes. The increasing use of multi-gene panel testing has led to the identification of many *TP53* variants in cancer patients outside LFS and LFL spectrum. Further, the majority of variants identified are missense alterations, complicating assessment of their clinical relevance. Methods used commonly to evaluate variants include qualitative criteria such as the ACMG guidelines, or statistical quantitative methods such as multifactorial likelihood analysis. Using a reference set of pathogenic (n=352) and benign (n=166) *TP53* missense variants defined by both clinical and functional evidence, this work aimed to: (i) optimise performance of the Align-GVGD missense prediction tool by modification of the existing multi-sequence alignment datasets; (ii) compare the performance of different bioinformatic prediction methods to address ACMG computational rules (PP3, BP4); (iii) investigate use of control datasets (gnomAD) to address ACMG population rules (BS1, BA1); and (iv) investigate other information sources that could be used as evidence in quantitative modelling (duplicated amino acid residues (ACMG PM5), somatic to germline ratio). Increasing the average number of substitutions per position in the multi-sequence alignments used by Align-GVGD improved performance, with ROC analysis AUC 0.909 for the new optimized alignment compared to 0.891-0.900 for existing alignments. The Matthews correlation coefficient (MCC) was then used to compare the prediction quality for several tools: optimized Align-GVGD, SIFT, Polyphen2, and the ensemble predictor REVEL. Prediction was more reliable when using a combination of tools (MCC = 0.83 for optimized Align-GVGD, SIFT, and Polyphen2; MCC = 0.84 for optimized Align-GVGD and REVEL with score cut-off of 0.7) than when using any tool alone (MCC = 0.55 – 0.77). Regarding population data, the gnomAD dataset included 699 unique *TP53* variants. The highest population-specific frequency (excluding Ashkenazi Jewish) of a single likely pathogenic variant of any effect type was 9.75e-5 for the missense variant LRG_321t1:c.848G>A (p.Arg283His). This observation supports using frequency cut-offs of 0.05% and 0.1% as ACMG strong and stand-alone evidence against pathogenicity, respectively. For the designated reference sets, co-location of missense substitutions was observed for 88.1% of the pathogenic variants and 58.4% of the benign variants, with only 16 unique amino acid residues shared between both sets. This observation indicates that finding a new variant in a known pathogenic residue could be used as qualitative evidence for pathogenicity (ACMG PM5), and has potential to be included as quantitative evidence in multifactorial modelling. The somatic to germline count ratio in the IARC *TP53* database was highly correlated for pathogenic (R^2 = 0.8044), but not benign (R^2 = 0.0706) reference set variants, and has demonstrated value for use in quantitative modelling (Paul James, submitted). In conclusion, we provide evidence to define gene/syndrome-specific data sources suitable for *TP53* variant classification using ACMG criteria, and/or quantitative statistical modelling.

POS025 - M

Variant detection, interpretation and reporting using an end-to-end analysis pipeline

Ruth Burton[1], Shyamal Dilhan Weeraratne[2] and Julie Deschênes[2]

[1]QIAGEN Advanced Genomics, Manchester, UK, [2]QIAGEN, Waltham, USA

Historically, detection, interpretation and reporting of variants has been a labour intensive and time-consuming process. A range of different software applications have been required to detect and filter variants and information from a variety of sources has been needed to accurately classify each variant. This process can be further complicated if there is a need to provide up to date information on relevant treatments and clinical trials. The final step of collating and summarising findings in a clinical report completes the process.
Using datasets from The Cancer Genome Atlas (TCGA) we present an end-to-end workflow starting with raw, unaligned reads from whole exome sequencing (WES) and ending with a comprehensive report describing clinically relevant variants. The first stage of the data analysis pipeline - to align the next generation sequencing (NGS) reads and detect variants was performed using Biomedical Genomics Workbench (BxWB). The list of variants was then refined using Ingenuity Variant Analysis (IVA). This filtering step, critical for WES, formed an integral, seamless part of the analysis pipeline in BxWB. The final steps of interpretation and report generation were performed by transferring the filtered variant list from IVA into QIAGEN Clinical Insight (QCI™) Interpret software and the automatic application of the American College of Medical Genetics (ACMG) Guidelines [1].
The combination of these three powerful software packages in a single pipeline were shown to reliably identify and classify clinically relevant variants. We were able to detect and classify disease causing variants in two colorectal cancer datasets including the KRAS variant p.G12R and the TP53 variant p.R273H. In addition using four breast cancer datasets we detected disease causing variants in BRCA2, PIK2CA, TP53 and MAP2K1. We were able to contextually associate these clinically significant variants with relevant treatment information and clinical trial recommendations to support patient care.

1. Journal article - Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody W, Hegde M, Lyon E, Spector E, VoelkerdingK, Rehm H (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine 17:405-423 doi: 10.1038/gim.2015.30

POS026 - T

Creating new national variation databases and their regulatory environments in Japan

Natsuko Yamamoto and Kazuto Kato

Department of Biomedical Ethics and Public Policy, Graduate School of Medicine, Osaka University, Japan

In 2016, with funding from the Japan Agency for Medical Research and Development (AMED), Japan started new programs focused on working towards connecting existing variation databases and enhancing the collection of human genomic variation and disease phenotypes at the national level.

At the same time, efforts in data protection and privacy preservation have become more vital all around the world. In Japan, for instance, genomic and clinical information and

the information concerning health have now been put under legal protection for the first time ever, thanks to a recent amendment to the Personal Information Protection (PIP) Laws in 2015 and 2016.
The authors are working as members of the Ethical, Legal and Social Issues (ELSI) task force as part of the AMED database programs. For this poster presentation, we will focus mainly on two current challenges facing the improvement of variation databases. The first one is how to collect the existing data, or legacy data. Many doctors, hospitals and research institutes have accumulated these kinds of data, and to make this data available via the national databases, we are making guidelines which adheres to the amended PIP Laws and other ethical guidelines for registrants of the database.

The second challenge we will focus on is global data sharing. The sharing and gathering of data is important for precise understanding of the relation between genomic variants and diseases. At the moment, one of the biggest hurdles in this task is the variety of regulations in different countries. For this poster presentation, we will present related Japanese laws and national rules, and would like to discuss how to promote international data sharing.

POS027 - M

A New National Initiative and Data Sharing Approach for Genomic Medicine in Japan

Jusaku Minari and Kazuto Kato

Department of Biomedical Ethics and Public Policy, Osaka University, Japan

Recently, leading programs such as the Precision Medicine Initiative in the US and Genomics England in the UK have promoted the implementation of genomic medicine. In addition to these programs in the US and the UK, several countries including Australia, France and China have also initiated government-led approaches for genomic medicine. These national-level initiatives aim to demonstrate the complex relationship between genetics, lifestyle and environmental factors regarding the development of disease, aiming for individualized diagnosis, prevention or treatment of disease.

In Japan, a new policy movement for genomic medicine has recently began. In 2012, Japan established the Headquarters of Healthcare Policy (HHP) to promote healthcare innovation. The committee in charge of 'genomic medicine', positioned under the HHP, has released a new policy for the achievement of genomic medicine, through showing the approach of research development including 'data sharing' as well as of improving the environment for genomic medicine. Regarding this research development, a new governmental funding agency, the Japan Agency for Medical Research and Development (AMED), has been newly established and strongly promotes research projects for genomic medicine, where a sharing policy of genomic information was created for the first time by a Japanese funding agency. In addition to these movements, the Act on the Protection of Personal Information was amended in 2015, which led to the revisions of governmental regulations regarding genomic information protection.

In this study, we have analyzed features of current policy movements and research developmental initiatives in Japan with regard to data sharing. We conclude that there are five key factors on genomic medicine initiatives in Japan; (1) the important role that the AMED data sharing policy holds, (2) the cooperation of several research projects with data sharing through funding agencies, (3) recognition of global data sharing initiatives through Global Alliance for Global Health, (4) main changes of governmental regulations and (5) an emphasis on ethical and legal

challenges in sharing genomic data both nationally and globally.
This study represents a valuable opportunity to share experiences regarding policy developments and research promotions involving data sharing in Japan with other countries, and to explore and develop new approaches for global collaboration in genomic medicine.

POS028 - T

Interpreting the functional effects of variants with integrated platforms from UniProt

Andrew Nightingale[1,2], Maria Martin[1,2], and The UniProt Consortium[1,2,3]

1 *EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK*
2 SIB Swiss Institute of Bioinformatics, Geneva, Switzerland
3 Protein Information Resource, Georgetown University, Washington DC & University of Delaware, USA

Precision medicine requires the detailed interpretation of genetic mutations and their consequences. However, almost all major computational methods treat genomic and the wealth of data from protein functional and structural studies independently. UniProt [1] is addressing this issue with new platforms for the interpretation of genetic variants in conjunction with the rich functional knowledge in the UniProt Knowledgebase (UniProtKB) and enriched per residue structural annotation of genetic variants.

Functional and structural protein annotations can be integrated into genome browsers with UniProt's *Genome Annotation Tracks*. We have mapped protein annotations in the human proteome to the GRCh38 assembly of the human genome. Thirty-three structural and functional annotations are currently provided including: enzyme active sites, modified residues, protein binding domains, protein variations etc; with the UniProt complete set of human proteome sequences, including isoforms, for comparison to predicted transcripts. These annotations and expressed proteome sequences are available as individual text BED files and are also bundled into a public track hub for track hub enabled genome browsers such as the Ensembl [2] and UCSC [3] genome browsers and can be accessed directly from the genome track hubs [4] registry (https://trackhubregistry.org/).

Genomic scientists often find protein structure and function data difficult to interpret; the *Protein Variant Effect Interpreter* (PepVEP) is an integration of the genomic, protein and structural information of a genetic mutation for interpreting the effect of the variant. This resource provides per residue functional information by integrating publicly available information and services from EMBL-EBI, such as, the Ensembl Variant Effect Predictor (VEP) [5] (Variation annotation), UniProt functional residue annotation (Protein function), and PDBe's structural residue annotation (Structure function). This web resource provides an integrated intuitive visualization illustrating the contributing information from individual resources for a comprehensive overview of the biological consequences of a genetic variant.

To illustrate how the *Genome Annotation Tracks* and PepVEP resources provide new and unique opportunities for biomedical research we examine biological examples in disease related genes and proteins illustrating how they can be utilized for the functional interpretation of variants. Thus, showing how they can help scientists to rapidly comprehend complex processes in biology.

1. The UniProt Consortium., 2017, UniProt: the universal protein knowledgebase *Nucleic Acids Res*. **45**: D158-D169.
2. Bronwen Aken, Premanand Achuthan, Wasiu Akanni, Ridwan Amode, Friederike Bernsdorff, Jyothish Bhai, Konstantinos Billis, Denise Carvalho-Silva, Carla

Cummins, Peter Clapham, *et al*. **Ensembl 2017,** *Nucleic Acids Research*, doi: 10.1093/nar/gkw1104
3. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002 Jun;12(6): 996-1006.
4. Raney BJ, et al. 2014. Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics* **30**: 1003-1005.
5. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P and Cunningham F.
The Ensembl Variant Effect Predictor *Genome Biology* 17:122(2016)

POS029 - M

In silico Prediction of Deleterious SNPs in Endocytosis Genes Implicated in Alzheimer's disease

HanJieh Tey and ChongHan Ng

Faculty of Information Science and Technology, Multimedia University, 75450 Bukit Beruang, Melaka, Malaysia
Email: chng@mmu.edu.my

From genome wide association studies (GWAS) on Alzheimer's disease (AD), it has been shown that many synonymous mutations (sSNP) and non-synonymous mutations (nsSNP) of genes from different pathways may affect the disease risk. One of the pathways associated with AD is endocytosis, in which mutations of endocytosis genes may disrupt Aβ clearance in the brain. This study is aimed to predict the impact of both sSNPs and non-synonymous missense mutations on three previously identified endocytosis genes, namely *PICALM*, *SYNJ1* and *SH3KBP1* on Alzheimer's disease by using computational methods. All the SNPs information was firstly retrieved from Ensembl genome database. Subsequently, PredictSNP2, which integrated 5 different prediction algorithms, was used to assess the deleteriousness of sSNPs to these genes in terms of evolutionary conservation, regulatory and transcription-binding motifs. From the result, a total of 8 sSNPs from both *PICALM* and *SYNJ1* genes were predicted to be deleterious. While for non-synonymous missense mutation, annotation tools including SIFT, PolyPhen-2, Mutation Assessor, I-Mutant Suite and SNPs&GO were used to predict the protein structural and functional impact of these mutations by using different methods and algorithms. There are 4, 12 and 1 missense mutations from *PICALM*, *SYNJ1* and *SH3KBP1* genes, respectively, predicted to be deleterious to the proteins structure and function. Interestingly, all the identified deleterious mutations were rare mutation (frequency < 1%), supporting the common disease rare variant (CDRV) hypothesis, in which rare mutation could be the major factor in common disease susceptibility. Besides that, most of the deleterious mutations fall on the known functional region of the protein. Computational analysis result showed that wild type and the substituted amino acids were different in terms of hydrophobicity, polarity and charges. Changing in hydrophobicity may cause loss of hydrophobic interaction while charges differences may lead to protein folding problem and loss of interaction with other molecules. Hence, the prediction analysis of sSNPs and nsSNPs in human endocytosis genes would be useful for further functional analysis.

Acknowledgement: This study was financially supported by Fundamental Research Grant Scheme (FRGS), Jabatan Pengajian Tinggi, Kementerian Pendidikan Malaysia.

POS030 - T

Application of fibre-FISH in the characterization of complex structural rearrangement and variation: an overview

Fengtang Yang

Molecular Cytogenetics Core Facility, Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom

The application of array-CGH and whole-genome sequencing has led the discovery of many genomic loci with highly variable structure, including copy number variants generated by duplications and deletions, as well as copy number neutral variants created by inversion and transposition. Nevertheless, characterisation of complex structural variation (SV), which is enriched with tandem duplications of low copy repeats and multi-allelic gene families, still represents a challenge to current high-throughput genomic methologies such as array-CGH, quantitative-PCR and next-generation sequencing. Fluorescence in-situ hybridisation (FISH), in particular fibre-FISH, remains the current method of choice for tackling such a challenge, since it enables direct visualisation of complex SVs and determination of SV haplotypes in a single hybridisation. In the past decade, our group have been involved in the validation and characterisation of a variety of complex genomic structures and SVs in human, mouse, zebrafish and pig by multi-colour fibre-FISH, using combed single-molecule DNA fibres. In this presentation, I will review the contribution of fibre-FISH to the further improvement of genome assemblies and to the determination of SV haplotypes of polymorphic SVs. I will demonstrate that by the judicial use of probes generated by long-range PCR and sequenced genomic clones (BACs and fosmids), both simple and complex genomic rearrangements as well as SVs can be resolved using fibre-FISH, and that fibre-FISH remains one of the most accurate methodologies for typing multi-allelic SVs.

POS031 - M

Overcoming the diagnostic challenges in neurological disorders: the role of next-generation sequencing

Sofia Gouveia[1]; Ana Fernández-Marmiesse[1], Francisco Laranjeira[2], Iria Roca[1], José Cocho[1], Daisy Castiñeiras[1], José Maria Fraga[1], Maria Luz Couce[1]

1:Unidad de Diagnóstico y Tratamiento de Enfermedades Metabólicas Congénitas, Hospital Clínico Universitario de Santiago de Compostela, Spain
2: Unidade de Bioquímica Genética, Centro de Genética Médica Jacinto Magalhães, Centro Hospitalar do Porto, Portugal

Purpose: Rare genetic diseases affect at least 1 in 50 individuals. There are an estimated 6000-7000 of these diseases, at least half of which are characterized by dysfunction of the central or peripheral nervous system. These clinical presentations involve the full range of the neuroaxis and include monogenic forms of brain malformations, ataxias, encephalopathies, myopathies and muscular dystrophies, neuropathies, movement disorders, metabolism defects, epilepsies, ciliopathies and dementias. Identifying the mutation(s) responsible for any of these varied clinical presentations can be extremely challenging. The chief difficulty resides in the genetically heterogeneous nature of neurological disease. The diagnostic process is often long and complex, with most patients undergoing multiple invasive and costly investigations without a conclusive molecular diagnosis ever being reached. The arrival of next-generation sequencing (NGS) coupled with advanced bioinformatics processing is changing the face of rare disease diagnosis by

offering faster, less expensive, and higher-resolution genetic testing. Although it is becoming increasingly more common for clinicians to use genomic data in their daily work for disease prevention, diagnosis, and treatment, the process of integrating genomic data into the practice of medicine has been a slow and challenging one. Genomic medicine programs are currently under way at several academic medical centers and large integrated health systems, but it is challenging to identify which genomic applications have robust evidence supporting their use in the clinic to improve patient outcomes.

Methodology: The present work is centered on a 4-year (2013-2016) cohort study that enrolled over 600 patients who were investigated for putative neurologic and/or metabolic conditions. We present a comprehensive description of the results from the application of an NGS-based platform (NeuroMeGen) to the diagnostic workflow of these patients. The platform consists of several customized gene panels optimized for different groups of disorders clustered by clinical and/or biochemical overlapping.

Results: We assess the power of the NeuroMeGen tool and look at its overall and panel-specific diagnostic yield (the epilepsy and neuromuscular panels offered the best results) and the involvement of de novo mutations and copy number variations in the mutational load of these diseases.

Conclusions: Finally, we discuss the possible causes of non-diagnosis and review the challenges and knowledge gaps identified during the development of NeuroMeGen. This work represents the first broad-scale approach to the implementation of genomic evaluation for the public health system in the Iberian Peninsula.

POS032 - T

Next-generation sequencing meets splicing: a multiplexed minigene splicing assay for exonic variants of OPN1LW and OPN1MW

Elena Buena-Atienza[1], Marius Codrea[2], Sven Poths[3], Marc Sturm[3], Karin Schaefferhoff[3], Sven Nahnsen[2], Bernd Wissinger[1]

[1] Institute for Ophthalmic Research, Centre for Ophthalmology, Tuebingen, Germany.
[2] Quantitative Biology Center, University of Tuebingen, Tuebingen, Germany.
[3] Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Tuebingen, Germany.

Purpose: Interpretation and prediction of the consequences caused by DNA variants in medical genetics still remain challenging. Whereas variants at canonical splice sites are correlated with splicing defects, functional classifying criteria for internal exonic variants are lacking. Experimental validation is limited to the availability of patient RNA material. Alternatively, minigenes have proved highly reliable [1]. Expressed in human cone photoreceptors, *OPN1LW* and *OPN1MW*, the opsin 1, long and medium wave sensitive genes are subject to frequent unequal homologous recombination and gene conversion events which may result in new haplotypes of uncertain relevance. Previously, we applied minigenes and showed that certain exon 3 haplotypes – defined by combinations of the following *per se* common coding SNPs: c.(453A > G; 457A > C; 465C > G; 511G > A; 513G > T; 521C> T; 532A > G; 538T > G) – induce exon skipping and are associated with X-linked cone dysfunction disorders [2]. However, systematic testing using minigenes is laborious. We thus introduce a parallelized minigene assay to assess the impact of hundreds of variants and haplotypes as

exemplified by *OPN1LW/MW* exon 3 haplotypes.

Methodology: We performed barcoding and site-directed mutagenesis of *OPN1LW/MW* exon 3 of the reference minigene [2]. Libraries were cloned by long homology overlaps (In-Fusion Cloning, IFC). The resulting "Input Library" was sequenced with PacBio. Reads were analysed using custom R scripts. HEK293T cells were transfected with the Input Library. Following reverse-transcription of target RNA, the cDNA pool was amplified to generate the "Output Library" for next-generation sequencing and quantification of correctly and aberrantly spliced transcripts.

Results: Sanger sequencing of clones showed that IFC efficiency was >90%. PacBio long reads enabled to phase each exon 3 haplotype with its cognate barcode. We successfully generated a combinatorial library of 256 possible haplotypes. 97% of the targeted haplotypes were tagged with at least 1 unique barcode and with 6 on average. Sanger-sequenced clones from the Output Library indicated that haplotypes inducing aberrant transcripts (i.e. lacking exon 3) are barcode-traceable. Deep sequencing of the Output Library is underway.

Conclusion: We present a minigene-based strategy for testing of hundreds of variants in parallel and which may be adapted for various exons and genes. This approach may also aid to decipher the grammar of exonic *cis*-regulatory elements of splicing.
This work was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration [grant 317472], and by BCM Families Foundation.

References

1. Steffensen AY, Dandanell M, Jønson L et al (2014). Functional characterization of BRCA1 gene variants by mini-gene splicing assay. Eur J Hum Gen 22: 1362–1368.
2. Buena-Atienza E, Rüther K, Baumann B, et al (2016). *De novo* intrachromosomal gene conversion from *OPN1MW* to *OPN1LW* in the male germline results in Blue Cone Monochromacy. Sci Rep 6: 28253.

POS033 - M

Utility of trio exome sequencing as a first-line diagnostic test for neurodevelopmental disorders

Ana Fernández-Marmiesse[1], Maria Soledad López-García[2], Begoña de Azua[2], Iria Roca[1], Sofia Gouveia[1], Maria Luz Couce[1].

1. Hospital Clinico Universitário de Santiago de Compostela, Santiago De Compostela, España
2. Hospital Son Llátzer, Palma De Mallorca, España

Purpose: 8 years-old patient, studied by psychomotor delay and myopathy. Asociated with weight loss, motor oral dysfunction, joint hypermobility. Performed studies: CPK, normal LDH, onset MRI normal, later cerebellar atrophy. Neurophysiological exploration with suggestive findings of a diffuse myogenic affection. Muscular biopsy compatible with centronuclear myopathy. WISC IV: light mental retardation. The analysis of the neuromuscular genetic panel was not conclusive. Identification of rare variants in the patient's exome which could be implied in her disorder.

Methodology: Simultaneous sequencing of the codifying regions (exons including binding regions exon-intron) in genes of the human genome using Next-Generation Sequencing technology consistent with in-solution hybridization enrichment with a Sure Select XT kit from Agilent specific for Whole-Exome analysis (SureSelectXT Human All exon v6) and posterior sequencing in a HiSeq 2000 from Illumina. The bioinformatic programs

used for the primary and secondary analysis were: RTA 1.8.70 (Illumina), FastQC v0.10.1, BWA v0.7.5, VarScan, SAMtools v0.1.19 and Annovar Nov2011. The variant filter and priorization was performed with in-house bioinformatic programs.

Samples: genomic DNA from the patient and her parents.

Results: This WES assay achieved at least 10X coverage in 96% of the analyzed regions for the trio samples. It was found 60.928 variants. The filter for exonic or *splicing* variants with a frequency <0.01 in 1000G database, decreased this number to 3.001. These variants were then filtered with the data of the patient's parents assuming two hypothesis: *de novo* or recessive. Through the first hypothesis one variant was detected: c.C1024T in heterozygous state in exon 8 of *CTBP1* gene (C-TERMINAL-BINDING PROTEIN 1; reference sequence NM_001328), which leads to an amino acid change, p.Arg331Trp. This gene was not previously associated to any pathology in OMIM, but its search in PubMed revealed an *in press* publication of Beck et al. where they described exactly the same mutation in four patients with similar clinical signs.
Conclusions: This work highlights the importance of using WES as a first-line diagnostic tool in the paediatric neurodevelopment disorders. The high importance of sharing findings of new genes associated with a specific phenotype within the international scientific community. In our case, if we did not find that bibliographic reference we would need to perform functional tests to prove the implication of this mutation in the patient's phenotype.

POS034 - T

Massive parallel sequencing for universal indications: diagnostic variant detection in a regional genetic center

Elisabeth Maurer, Martina Witsch-Baumgartner, Johannes Zschocke

Medical University Innsbruck, Division of Human Genetics, Innsbruck, Austria

Molecular diagnostics has shifted from institutes with special expertise for individual indications to a unified diagnostic pipeline in regional centres. The latter approach poses special challenges with regard to a wide range of referral reasons, the possibility and need to obtain detailed phenotype descriptions, adequate patient consent, optimal analysis and interpretation pipeline, and counseling of patients. Here, we present a comprehensive strategy that takes all these aspects into consideration. Our approach starts with the individual patients seen in the genetic clinic or by regional partners to adequately document the phenotype and obtain consent. Based on the clinical needs and the number and spectrum of target genes, one of a limited set of library preparations is chosen for massive parallel sequencing. Enrichment sets include commercially available organ-specific panels, a "house panel" used for frequently requested genes, clinical exomes (OMIM disease genes) or full exomes. Analyses entail deletion/duplication assessment using several tested programs. In most circumstances, target genes only are examined and non-relevant genes on the panels remain blinded to minimize challenge of interpretation and preclude unconsented results. In a multi-step process, additional genes may be analyzed at a later stage without additional laboratory costs. Massive parallel sequencing is complemented by all relevant diagnostic methods including routine transcript sequencing, fragment analyses, Southern blot, DNA array, FISH, etc. This mode of handling in

our laboratory regarding diagnostics of inherited diseases permits to deliver high quality reports and answers specific questions of referring physicians in a regional setting; examples of successful application of the pipeline will be presented.

POS035 - M

Targeted Next- Generation Sequencing For Molecular Diagnosis Of Ichthyosis

Uxia Esperón-Moldes[1], Manuel Ginarte[2], Laura Rodríguez-Pazos[3], Laura Fachal[1], Ana Vega[1]
[1]Fundación Pública Galega de Medicina Xenómica-SERGAS, CIBERER, IDIS, Santiago de Compostela, España.

[2]Servicio de Dermatología del Complexo Hospitalario Universitario de Santiago, Facultad de Medicina, Santiago de Compostela, España.
[3]Servicio de Dermatología del Complejo Hospitalario Universitario de Vigo, Vigo, España.

Ichthyosis is a group of skin genetic disorders that can follow a recessive or dominant inheritance pattern. Manifestations can be restricted to the skin or associated with extracutaneous symptoms. ARCI (Autosomal Recessive Congenital Ichthyosis) is a rare non syndromic type of ichthyosis. Their diagnosis is based on clinical manifestations which is often a complex task due to their phenotypic variety. They also present genetic heterogeneity posing a significant challengefor the molecular diagnosis. Advances in sequencing technologies can reduce this complexity in a cost-effective manner simplifying the detection of mutations in this group of patients [1, 2].

Purpose: The aim of this study is to evaluate the utility of a targeted next-generation sequencing panel in the differential diagnosis of ARCI in Spanish patients.

Methodology: Fifty six patients with a clinical suspicion of ARCI were recruited from different dermatology services of Spanish hospitals and the national patient organization for ichthyosis (ASIC). Their clinical information, consent forms and blood samples were provided. Targeted resequencing was performed on their DNA samples using SOLiD 5500xl or Ion Proton Platforms. A 40 gene panel was designed for ichthyoses genetic diagnosis. Variants detected were validatedbySanger sequencing,or MLPA(Multiplex Ligation-dependent Probe Amplification)for large deletions.

Results/Discussion: Genetic analysis revealed mutations in twelve different genes. Missense variants were the most prevalent (51%) type of mutations. We identified the causative mutation in 44 patients (79%).Mutation status was analyzed in first degree relatives when possible. All patients were initially diagnosed with ARCI. However, we found mutations in genes related with other ichthyoses such as X-linked ichthyosis, epidermolytic ichthyosis, ichthyosis prematurity syndrome and ARCI with hypotrichosis. Thus, next generation sequencing led to a reclassification of the clinical diagnosis for up to 14% of the patients with identified mutations.

Conclusions: Our findings demonstrate the utility of our panel for the differential diagnosis of different types of ichthyoses. This study also highlights the importance of next generation sequencing panels in the molecular diagnosis and genetic characterizationof patients with ARCI and other types of ichthyoses.

Acknowledgements: We are grateful to patients, families and dermatologists for their cooperation.

Founding sources: Ramón Areces Foundation project awarded to A.V.

References:

1. Fischer J (2009) Autosomal recessive congenital ichthyosis. J Invest Dermatol 129:1319-21. doi: 10.1038/jid.2009.57.
2. Diociaiuti A, El Hachem M, Pisaneschi E, Giancristoforo S, Genovese S, Sirleto P, Boldrini R, Angioni A (2016) Role of molecular testing in the multidisciplinary diagnostic approach of ichthyosis. Orphanet J Rare Dis 11: 4. doi: 10.1186/s13023-016-0384-4.

POS036 - T

Cycling Temperature Capillary Electrophoresis: A quantitative, fast and inexpensive method to detect mutations in mixed populations of human mitochondrial DNA

Paulo Refinetti[a], Stephan Morgenthaler[a], Per O. Ekstrøm[b]

[a] *École Polytechnique Fédérale de Lausanne, EPFL FSB STAP, Station 8, Lausanne, Switzerland* [b] *Department of Tumor Biology, Norwegian Radiumhospital, Oslo, Norway*

Cycling temperature capillary electrophoresis has been optimised for mutation detection in 76% of the mitochondrial genome. The method was tested on a mixed sample and compared to mutation detection by next generation sequencing. Out of 152 fragments 90 were concordant, 51 discordant and in 11 were semi-concordant. Dilution experiments show that cycling capillary electrophoresis has a detection limit of 1-3%. The detection limit of routine next generation sequencing was in the ranges of 15 to 30%. Cycling temperature capillary electrophoresis detect and accurate quantify mutations at a fraction of the cost and time required to perform a next generation sequencing analysis.

POS037 - M

Germline mutations in childhood cancer patients suspected of genetic predisposition to cancer – a retrospective analysis

Dianne E Sylvester 1, Yuyan Chen 1, Robyn Jamieson 2, Luciano Dalla Pozza 3, and Jennifer A Byrne 1.

1. Children's Cancer Research Unit, Kids Research Institute, Discipline of Child & Adolescent Health, University of Sydney, Westmead, NSW, Australia.
2. Eye and Developmental Genetics Research, Children's Medical Research Institute, Discipline of Child & Adolescent Health and Discipline of Genetic Medicine, University of Sydney, Westmead, NSW, Australia.
3. The Cancer Centre for Children, The Children's Hospital at Westmead, NSW, Australia.

Purpose: The proportion of childhood cancer patients with predisposing constitutive genetic mutation(s) is not clearly defined. In many treating centres the majority of childhood cancer patients are not offered germline testing even if a hereditary component is suspected. This is partly due to the cost and time-consuming nature of gene-targeted diagnostic tests. With the growing availability of massively parallel sequencing technologies, it is now possible to offer families a single sequencing test for a panel of cancer predisposition genes. We studied a retrospectively identified childhood cancer patient cohort using whole exome sequencing to identify pathogenic germline mutations, whilst also considering novel variants which may predispose to disease.

Methodology: The inclusion criteria were childhood cancer patients diagnosed at the Children's Hospital at Westmead (1997-2015) suspected of genetic predisposition to cancer

for whom frozen blood samples were available for study, specifically siblings developing cancer, patients developing multiple cancers, and/or childhood cancer in association with either a genetic syndrome and/or significant family history of malignancy. DNA was extracted from frozen blood samples and underwent whole exome sequencing using Agilent SureSelect targeted enrichment and sequencing on an Illumina HiSeq2000 (Macrogen, Korea).

Results: The data (approximately 90,000 variants per exome) were annotated with ANNOVAR [1] and filtered for very rare (less than 0.1% Exome Aggregation Consortium population) exonic variants in genes associated with cancer predisposition, somatic mutations in cancer and/or DNA repair (n=1047 genes). Each patient had a mean of 20 variants (range 7-62) that met the criteria. Variants were interpreted for pathogenicity using the American College of Medical Genetics and Genomics guidelines. To date, 10/39 (26%) patients carried a pathogenic/likely pathogenic germline mutation in a known cancer predisposition gene.

Conclusion: In this study of retrospectively identified childhood cancer patients suspected of genetic predisposition, over a quarter carried germline mutation(s) either known or very likely to predispose to cancer. As also previously observed [2,3], the landscape of germline mutations in childhood cancer patients is more extensive than traditionally assumed. Further to this, with additional investigations, some variants of uncertain significance may also be shown to contribute to cancer predisposition in childhood.

1. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data. Nucleic Acids Research. doi:10.1093/nar/gkq603
2. Zhang, J., et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. N Engl J Med. doi:10.1056/NEJMc1600338
3. Parsons, et al. Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. JAMA Oncol. doi:10.1001/jamaoncol.2015.5699

POS038 - T

Dissecting the molecular basis of epileptic disorders in the Iberian Peninsula

Sofia Gouveia[1]; Ana Fernández-Marmiesse[1], Francisco Laranjeira[2], Iria Roca[1], José Cocho[1], Daisy Castiñeiras[1], José Maria Fraga[1], Maria Luz Couce[1]

1: Unidad de Diagnóstico y Tratamiento de Enfermedades Metabólicas Congénitas, Hospital Clínico Universitario de Santiago de Compostela, Spain
2: Unidade de Bioquímica Genética, Centro de Genética Médica Jacinto Magalhães, Centro Hospitalar do Porto, Portugal

Purpose: Epilepsy is one of the most common neurological conditions, with a prevalence of 1%. The identification of genes associated with epilepsy has presented major challenges for many years due to 1) a highly variable inter and intra-familial expressivity of phenotypes 2) a high frequency of phenocopies, and 3) enormous genetic heterogeneity. Recent technological breakthroughs such as aCGH and NGS have led to the identification of an increasing number of genomic regions and genes responsible for epileptic disorders (in particular epileptic encephalopathies), uncovering a wide spectrum of pathophysiological mechanisms. *De novo* dominant mutations are frequently identified, and somatic mosaicism and recessive disorders are also seen. Gene discovery provides the basis for neurobiological insights and frequently reveals convergence of mechanistic pathways. Findings such as these underpin the development of targeted therapies, which are essential for improving

the outcome of these devastating disorders. This work presents a comprehensive description of the results of the application of an NGS-based platform (NeuroMeGen) to the diagnostic workflow of 135 patients with epileptic disorders collected by leading neurology units in the Iberian Peninsula.

Methodology: The NeuroMeGen platform encompasses 1) a custom panel design updated periodically with new associated genes suggested by scientific publications; 2) a purpose-designed prioritization algorithm that takes into account specific population frequencies, mutation-susceptibility profiles for specific mutations, and different sensitivities for variant calling algorithms: and 3) an in-house developed CNV detection method (PattRec).

Results: The overall diagnosis rate for the molecular platform NeuroMeGen over the 4 years of our study is between 40-50%. The most frequently mutated genes identified in our population are *SCN1A*, *KCNQ2*, *CDKL5*, and *MECP2*. However, many patients have also been found to carry mutations in one of the other 186 genes currently included in NeuroMeGen. *De novo* mutations and CNVs constitute an important percentage of the genetic burden in the cases identified. Our results also support an important pathogenic role for mosaicism in epilepsy, since some patients have a pathogenic mutation also carried by one of their progenitors.

Conclusion: In sum, this work shows that the introduction of this molecular tool in the diagnostic protocol for epileptic disorders shortens time to diagnosis, decreases the number of tests and treatments required (most of which are unnecessary and even counterproductive), and guides genetic counseling for families who intend to have offspring.

POS039 - M

PGD counselling for variants of unknown significance

A. Orlova[1], Ya.Sofronova[1], N. Vetrova[1], E. Musatova[1,2], M. Skoblov[2], E.Pomerantseva[1]

1- Center of Genetics and Reproductive Medicine Genetico, Moscow, Russian Federation,
2 - FSBI "Research Centre for Medical Genetics", Moscow, Russian Federation

Purpose
Preimplantation genetic diagnosis (PGD) for monogenic disorders is a well-established laboratory procedure and from technical point of view PGD is possible for most of mutations. However decision on the feasibility of PGD could be difficult if pathogenicity of the mutation is uncertain. This uncertainty is typical when molecular diagnosis was established by NGS and mutation was classified as a variant of unknown significance (VOUS). In such cases, PGD laboratory has to conduct separate research and reach conclusion if there is enough evidence that mutation is pathogenic. Several methods could be useful, including bioinformatic analysis, screening of scientific databases and publications regarding gene in question, haplotyping and segregation analysis, and functional analysis.
We present a case of PGD for hereditary spherocytosis, where PGD counsellor suggested functional analysis of the mutant gene to confirm causative status of the mutation.

Methodology
Family *B* was seeking IVF because of infertility. PGD-counselling was suggested as partner had hereditary spherocytosis. Family wanted to prevent birth of an affected child after IVF and requested PGD. Only clinical diagnosis was known and no genetic testing of the partner was done before. As there are several genes linked to this condition, clinical

exome sequencing was chosen as a method of molecular study. In *SPTB* gene mutation c. 5798+2A>C was found and classified as VOUS. Other mutations in this gene are known to cause autosomal dominant form of hereditary spherocytosis. Functional analysis was performed on RNA, extracted from proband's blood . After RNA isolation and cDNA synthesis, expression analysis of target gene isoforms was made.
PGD system was developed both for direct mutation analysis by RFLP and for indirect linkage analysis by STR markers. Mutant haplotype was established by single sperm analysis. Chromosomal screening of healthy embryos was performed by aCGH on microarrays 24sure (Illumina).

Results
Functional analysis has shown that mutation leads to exon skipping therefore confirming pathogenicity of the mutation. One IVF cycle was conducted and 11 embryos were biopsied for PGD at 5th day of embryo development. PGD for *SPTB* mutation was performed and 4 embryos that inherited the mutation were excluded from further analysis. Remaining 7 embryos were screened for aneuploidy and 6 of them were confirmed to be normal.
There were several other cases where family requested PGD for VOUS. In family *V* two mutations in *CLCN7* gene chr16:1510924 C>T and chr16:1510924 C>T were regarded as pathogenic after bioinformatics analysis and study of segregation. In family *P*, however, PGD-lab could not confirm pathogenicity of the mutation arr[hg19] 5p13.3p13.2x3 and PGD was postponed until additional information on this VOUS becomes available.

Conclusions
These cases represent an example of a novel problem facing PGD laboratories – decision-making on VOUS. We believe that PGD laboratory has to be aware of potential risks associated with diagnosis of such mutations. However, PGD counselor might conduct additional research on the mutation and among other methods functional analysis of gene product is especially useful.

POS040 - T

Comprehensive variant analyses including whole genome sequencing in hereditary colorectal cancer syndromes

Anna Rohlin[1,2], Anders Kvist[2], Therese Törngren[2] Frida Eiengård[1], Ulf Lundstam[3], Theofanis Zagoras[1], Åke Borg[2], Jan Björk[4], and Margareta Nordling[1]

1. Department of Clinical Pathology and Genetics, Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden, 2. Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Medicon Village, Lund, Sweden, 3. Department of Surgery, Sahlgrenska Academy at University of Gothenburg, Sahlgrenska University Hospital/Östra, Gothenburg, Sweden, 4. The Swedish Polyposis Registry, Department of Medicine, Karolinska Institute, Stockholm, Sweden

High penetrant pathogenic variants in known hereditary colorectal cancer (CRC) predisposition genes explain 5-10% of cases. Clinical testing in hereditary CRC syndromes is usually performed based on phenotype of the different syndromes. However, overlapping phenotypes between CRC syndromes, including both polyposis and non-polyposis present diagnostic difficulties in the clinic. It is crucial that the patient get a correct molecular diagnosis that allows for adequate follow-up since the majority of the syndromes include predisposition for tumors also in other organs.

Analyses of SNPs, indels and CNVs were previously performed on approximately 100 individuals divided into clinical subtypes based on phenotype. The patients were

analyzed for pathogenic variants using a panel consisting of 19 high-risk CRC susceptibility genes including whole gene regions (1). In patients diagnosed with classical FAP, attenuated FAP, atypical FAP and non-polyposis subgroups, we identified pathogenic variants in *BMPR1A* and *SMAD4*. We also detected novel CNVs in upstream regions of *SMAD4*, *MSH3*, *CTNNB1* and one deletion in an intronic region of *CDH1*(1). Twenty-six of these patients have now been analyzed with whole genome sequencing (WGS). By using WGS novel SNPs, indels and CNVs were detected. One CNV was a novel complex CNV of the *SHROOM2* gene.

Using a comprehensive CRC gene panel we obtained increased detection frequency of pathogenic variants for CRC syndromes, which enable appropriate follow up of patients based on the clinical feature of each syndrome. Analysis by WGS identified additional variants and complex structural rearrangements.

1) Rohlin A, Rambech E, Kvist A, Törngren T, Eiengård F, Lundstam U, Zagoras T, Gebre-Medhin S, Borg Å, Björk J, Nilbert M and Nordling M. "Expanding the genotype-phenotype spectrum in hereditary colorectal cancer by gene panel testing", Fam Cancer, 2016, Sep 30 (Epub ahead of print)

POS041 - M

Evaluation Of Basic Massive Parallel Sequencing Parameters In Relation To True/False Positivity's Findings Of Rare Variants From Isolated Population With High Incidence Of Parkinsonism

1Radek Vodicka, 1Radek Vrtel, 1Kristyna Kolarikova, 1Martin Prochazka, 2Katerina Mensikova and 2Petr Kanovsky
1Department of Medical Genetics Faculty of Medicine and Dentistry, Palacky University, UniversityHospital, Olomouc, Czech Republic
2Department of Neurology Faculty of Medicine and Dentistry, Palacky University, UniversityHospital, Olomouc, Czech Republic

Introduction: Massive Parallel Sequencing (MPS) in 16 genes known to be associated with Parkinsonism including coding DNA, intron/exon boundaries and UTRs loci was used to find rare variants in 30 patients and 12 healthy controls from an isolated population of South-Eastern Moravia in the Czech Republic. Epidemiological data show significantly increased prevalence of Parkinsonism (2.9%).

Aim of the study: to evaluate true/false positivity ratio in relation to the basic MPS sequencing parameters (coverage, type of mutation – SNV/INDEL, percentage of rare variants in case of heterozygozity, +/- strand bias and length of homopolymers).

Method: Final filtered out rare variants were obtained from the Ion Torrent platform with workflow as following: Torrent Suite Base calling and BAM mapping, IonReporter Variant calling and rare variant filtering. True from false positivity findings were distinguished by Sanger confirmation sequencing.

Results: In total, there were found 36 rare variants (MAF< 1 %) from which 50 % of them were confirmed as true positive. Comparison between true and false positivity is displayed in table below:

Conclusion: In case of SNV, the probability of false positivity is 11.7 % while in INDEL the false positivity proportion is 84 %. Very interesting indicator of true positivity could be high correlation in strand biases of reference and rare variants in heterozygous findings.
This study was supported by grant MZ-NV15-32715A.

POS042 - T

Potentially pathogenic germline CHEK2 and NOTCH3 variants among multiple early-onset cancer families

Mev Dominguez-Valentin[1], Sigve Nakken[1,2], Helene Tubeuf[3,4], Daniel Vodak[1,2], Per Olaf Ekstrøm[1], Anke M. Nissen[5,6], Monika Morak[5,6], Elke Holinski-Feder[5,6], Alexandra Martins[3], Pål Møller[1,7-9], Eivind Hovig[1,2]

[1]Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway
[2]Department of Informatics, University of Oslo, Norway
[3]Inserm-U1245, UNIROUEN, Normandie Univ, Normandy Centre for Genomic and Personalized Medicine, Rouen, France
[4]Interactive Biosoftware, Rouen, France
[5]Medizinische Klinik und Poliklinik IV, Campus Innenstadt, Klinikum der Universität München, Ziemssenstr. 1, Munich, Germany
[6]MGZ—Medizinisch Genetisches Zentrum, Munich, Germany
[7]Institute of Cancer Genetics and Informatics, Oslo University Hospital, Oslo, Norway
[8]Department of Medical Genetics, Oslo University Hospital, Norway
[9]Department of Human Medicine, Universität Witten/Herdecke, Germany

Purpose: To study the role of other genes than *PTEN, TP53*, *BRCA1* and *BRCA2* associated to multiple early-onset cancer families including individuals having Cowden (CS) or Li-Fraumeni (LF) syndromes features but not meeting the full diagnostic criteria and early-onset breast cancer (BC) in Norwegian families, which may influence the biological and clinical characteristics of these families.

Methodology: The hereditary cancer biobank from the Norwegian Radium Hospital was used to identify multiple early-onset cancer families including individuals with Cowden-like syndrome (CSL), Li-Fraumeni-like syndrome (LFL) and early-onset BC diagnosis where no pathogenic variants in *PTEN, TP53* and *BRCA1/2* genes had been found by diagnostic DNA sequencing, respectively. Forty-four cancer susceptibility genes were selected and analyzed by our in-house designed TruSeq amplicon-based assay for targeted resequencing (TSCA v.1.5, Illumina, Palo Alto, CA). Protein and splicing-dedicated *in silico* analyses were performed for all unclassified single-nucleotide variants (SNVs) in order to predict their impact on splicing. Then, selected variants were experimentally analyzed by comparing the splicing pattern of reporter minigene constructs transfected into human cells.

Results: One individual carried a variant in *CHEK2* (c.319+2T>A), here considered as likely pathogenic given its position at the 5' splice site of exon 2. From the 5 SNVs tested in the splicing minigene assay, only *NOTCH3* c.14090C>T (p.Ser497Leu) showed a significant effect on RNA splicing, notably by inducing partial skipping of exon 9.

Conclusion: Among CSL, LFL and early-onset BC patients, multiple-gene sequencing identified a potentially pathogenic variant in *CHEK2* and another in *NOTCH3,* both affecting RNA splicing signals. Our study provides new information on variants on genetic loci that may affect the risk of developing cancer in these patients and their families, demonstrating that genes presently not routinely tested may be important for capturing predisposing for cancer development in these families.

Acknowledgements
We thank the families for their participation and contribution to this study. This work was supported by the Radium Hospital Foundation (Oslo, Norway), Helse Sør-Øst (Norway), the French Association Recherche contre le Cancer (ARC), the Groupement des Entreprises Françaises dans la Lutte contre le Cancer (Gefluc), the Association Nationale de

la Recherche et de la Technologie (ANRT, CIFRE PhD fellowship to H.T.) and by the OpenHealth Institute.

POS043 - M

Allelic drop-out is a common phenomenon in the PCR-based NGS and emphasizes the importance of cross-validation

Anna Bukaeva, Elena Zaklyazminskaya

Petrovsky Russian Research Centre of Surgery, Moscow, Russia.

Purpose. Capillary Sanger sequencing is widely used method for validation of the NGS results. To date, high-throughput semiconductor sequencing with AmpliSeq primer panels followed by validation of clinically significant variants by Sanger resequencing is a well-accepted way to perform DNA diagnostics on a set of genes in a short time with high cost-effectiveness. Nevertheless, as both AmpliSeq and Sanger are PCR-based methods, they have similar spectrum of the technical limitations. The purpose of this study is to demonstrate the allelic drop-out (ADO) as a common phenomenon that might decrease the accuracy of DNA diagnostics results.

Methodology. We had developed 3 custom gene panels using manufacturing sets of AmpliSeq oligoprimers for mutational screening. Panel “Long QT/Brugada syndrome” contains 234 primer pairs to encompass coding and canonic intronic areas of the 11 genes (about 30 kb); panel “TGF pathway” contains 342 primer pairs for the 14 genes (about 39 kb); and panel “DCM/ Desmosomal proteins” contains 521 primer pairs for the 16 genes (about 94 kb). All potentially significant variants and fragments with low coverage were re-analyzed with Sanger sequencing. Totally 160 patients were screened with at least one of these gene panels.

Results. We had found 5 ADO cases in the 5 DNA samples. One allele was dropped out using NGS (1/160, 0,6%) but then was discovered by control capillary resequencing. Four heterozygous SNPs (4/160, 2,5%) were revealed by IonTorrent sequencing in the control DNA samples. These samples were previously considered as “normal homozygous” after primary Sanger sequencing. So, we had detected at least one ADO case in 3.1% DNA samples using AmpliSeq-based NGS with following Sanger sequencing. All ADO cases were discovered occasionally because of loss of heterozygosity in the marker SNP located within the same PCR-product. We realize that in the vast majority of the amplicons we had no any marker SNP to visualize the number of alleles amplified. It seems that real scope of ADO might be much higher but we have no tool to estimate it.

Conclusion. All PCR-based methods have a risk of allelic drop-out. This risk is rising along with increasing the number of oligoprimers and the size of the DNA target. We assume that the best way to analyze extended fragment(s) of nucleic acid is the old “gold standard” of the double sequencing with two independent pairs of the oligos. It is time- and money-consuming method but it is the most accurate approach for the diagnostic needs.

This work was supported by Russian Science Foundation grant № 16-15-10421

POS044 - T

Challenges of interpreting sequence variants: experience of a molecular laboratory with a panel of epileptic encephalopathy

Susana Sousa[1,3]; Paulo Silva[1,3], Isabel Alonso[1,2,3]; Jorge Sequeiros[1,2,3,4]

[1.] Centro de Genética Preditiva e Preventiva - CGPP, Instituto de Biologia Molecular e Celular - IBMC, Universidade do Porto, Porto, Portugal.

[2.] UnIGENe, Instituto de Biologia Molecular e Celular - IBMC, Universidade do Porto, Porto, Portugal.

[3.] Instituto de Investigação e Inovação em Saúde - i3S, Universidade do Porto, Portugal.

[4.] Instituto de Ciências Biomédicas Abel Balazar, ICBAS, Universidade do Porto, Portugal.

Introduction: Sequencing technology has evolved rapidly with the advent of high-throughput next-generation sequencing. These advances have led to an explosion in gene discovery for many human disorders, including epilepsy. Epileptic encephalopathies (EE) are a phenotypically and genetically heterogeneous group of severe epilepsies accompanied by intellectual disability and other neurodevelopmental features. The increasing complexity led to a shift in genetic testing which has been accompanied by new challenges in variant interpretation.
CGPP uses the ACMG guidelines that recommend the use of specific standard terminology - “pathogenic,” “likely pathogenic,” “uncertain significance,” “likely benign,” and “benign” – to describe variants identified in genes that cause Mendelian disorders. Moreover, this recommendation describes the process for classifying variants into these five categories based on criteria using typical types of variant evidence (e.g., population frequencies, computational predictions, functional analysis, and segregation data).

Methods: Using a virtual gene panel (45 genes) for epileptic encephalopathies, sequencing was performed using Whole Exome Sequencing (WES) on an Illumina HiSeq 4000, and variants were identified using a custom pipeline, based on BWA for alignment (GRCh37), GATK HaplotypeCaller for variant calling and Ensembl VEP, GEMINI and Alamut for annotation.

Results: Our pipeline identified variants not previously described, which were described as “uncertain significance” in three different patients with epileptic encephalopathy.
Patient 1 – 3 variants were detected in heterozygosity; two different variants in *GRIN2A* gene – responsible for an autosomal dominant form of EE (c.1354G>A and c. 3925C>G) and one variant in *PCDH19* gene – an X-linked form (c.3235C>G). In this case, c. 3925C>G in *GRIN2A* gene is the variant that could, most probably, explain the phenotype.
Patient 2 – Two variants were detected in heterozygosity; one variant in *DEPDC5* gene – responsible for an autosomal dominant form (c.4418A>G) and other variant in *NRXN1* gene – an autosomal recessive form (c.479C>T). In this case, c.4418A>G in *DEPDC5* may be compatible with the patient clinical symptoms.
Patient 3 – 2 variants were detected in heterozygosity; one variant in *POLG* gene (c. 1550G>T); and other variant in *SCN1B* gene (c.566C>T). Both variants are of uncertain clinical significance and none can explain the patient phenotype.

Conclusion: The use of high-throughput sequencing technologies in diagnostic settings is becoming more prominent, although these data needs to be interpreted with great caution given the complexities outlined previously. It is critical to reach consensus between healthcare providers and clinical laboratories, so that both have a common understanding of how variants are classified and, thus, are able to provide a better service of patient counselling and management.

POS045 - M

Copy number variation differences between responders and non-responder to anti-TNF drugs in moderate-to-severe psoriasis

María C Ovejero-Benito[1], Ancor Sanz-García[2], Reolid A[3], Muñoz-Aceituno E[3], Mar Llamas-Velasco[3], Miriam Saiz Rodríguez[1], Teresa Cabaleiro[1], María Talegón[1], Esteban Daudén[3*], Francisco Abad-Santos[1,4*]

[1]Clinical Pharmacology Department, Hospital Universitario de la Princesa, Instituto Teófilo Hernando, Universidad Autónoma de Madrid (UAM), Instituto de Investigación Sanitaria la Princesa (IIS-IP), Madrid, Spain
[2]Clinical Neurophysiology Department, Hospital Universitario de la Princesa, Instituto de Investigación Sanitaria La Princesa (IIS-IP), Madrid, Spain
[3]Dermatology Department, Hospital Universitario de la Princesa, Instituto de Investigación Sanitaria La Princesa (IIS-IP), Madrid, Spain
[4]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Instituto de Salud Carlos III, Madrid, Spain
**: F Abad-Santos and E Daudén contributed equally to the manuscript*

Purpose - To identify copy number variations that could predict anti-TNF drugs response in moderate-to-severe psoriasis patients.

Methods- DNA was isolated from peripheral blood cells of 70 moderate-to-severe psoriasis patients treated with anti-TNF drugs. Effectiveness of anti-TNF agents was evaluated by the PASI (Psoriasis Area and Severity Index). Patients who achieved a 90% improvement over their baseline PASI (PASI90) were considered excellent responders to treatment (N=49) and patients who did not achieve a 75% improvement over baseline PASI (PASI75) were considered partial insufficiently/non-responders to treatment (N=21). Samples were bisulfite converted and hybridized in a 450K methylation microarray. Methylation data was corrected for cellular heterogeneity by estimating and adjusting for cell-type proportions in our blood-derived DNA samples. Methylation values were adjusted by the age and gender of the patients. Copy number alterations were detected using diverse R packages such as conumee, CopyNumber450k, DNAcopy and the ChAMP pipeline.

Results- Copy number alterations have been observed between partial insufficiently/non-responders. These copy number variations span through the diverse chromosomes and can be distributed in gain and loss of genes.

Conclusions- The participation of copy number alterations in the modulation of anti-TNF drugs response in moderate-to-severe psoriasis patients have not been studied so far. We have shown that copy numbers alterations may discriminate between excellent responders and partial insufficiently/non-responders to anti-TNF drugs in moderate-to-severe psoriasis patients.

POS046 - T

Identifying limiting factors for single nucleotide and copy number variant calling in a targeted sequencing panel

P. Cacheiro[1,2], B. Quintáns[1,2], J. Amigo[2,3], A. Ordóñez-Ugalde[1,2], A. Carracedo[2,3], M.J. Sobrido[1,2]

[1] Neurogenetics Group, Instituto de Investigación Sanitaria de Santiago (IDIS), [2]Grupo de Medicina Xenómica, CIBERER-U711, [3] Fundación Pública Galega de

Medicina Xenómica, Santiago de Compostela, Spain.

Purpose. Variant detection protocols for clinical applications of next generation sequencing (NGS) need evaluation and standardization. The human reference genome, against which the variants are called, contains both common and rare variants. The aim of this study was to analyze the performance of variant calling and copy number detection algorithms on an NGS panel for a rare condition. We also evaluated the implications of low frequency alleles present in the reference assembly, particularly in the case of monogenic recessive disorders.

Methodology. We evaluated the calling performance of a 30 gene panel sequenced in 83 patients with hereditary spastic paraplegia both for single nucleotide (SNV) and copy number (CNV) variants. The variant calls obtained with LifeScope, GATK UnifiedGenotyper and GATK HaplotypeCaller were compared with Sanger sequencing. Variant calling efficiency was evaluated for 187 (56 unique) SNVs and indels. Five multiexon deletions detected by multiple ligation probe assay were assessed with ExomeDepth, panelcn.MOPS and CNVPanelizer from the NGS panel data. Additionally, the 1000G phase 3 dataset was used to investigate those positions where the reference allele had a minor allele frequency ≤ 2%. These variants were then annotated and checked against HGMD, ClinVar and OMIM databases.

Results. 48/51 (94%) SNVs and 1/5 (20%) indels were consistently detected by all the calling algorithms. Two SNVs were not detected by any of the callers because of a rare reference allele, and one SNV in a low coverage region was only detected by two algorithms. Four indels were called with discrepancies in annotation and zygosity by each of the algorithms. Regarding CNVs, ExomeDepth detected 5/5 multi-exon deletions, panelcn.MOPs 4/5 and only 3/5 deletions were accurately detected by CNVPanelizer. The analysis of the reference assembly revealed that 14 low-frequency SNVs in the reference sequence were associated to disease in any of the three repositories, 7 of them in genes with an autosomal recessive mode of inheritance. Conversion to GRCh38 assembly showed that the reference allele was no longer the rare allele for 5/14 variants. We identified 96 additional variants with potential functional impact in recessive genes.

Conclusion. The calling efficiency of NGS algorithms for SNVs is influenced by variant type and coverage. NGS protocols need to account for ambiguities in indel calling. The presence of rare variants in the reference sequence should be taken into account, particularly for recessive disorders. CNV detection algorithms can be used to identify large deletions from NGS panel data for diagnostic applications, however sensitivity depends on coverage, selection of the reference set and deletion size. Incorporation of several variant callers in the NGS pipeline seems advisable to maximize variant detection efficiency.

Funding sources: This study was supported by Instituto de Salud Carlos III PI 13/01598, Fundación Teófilo Hernando, the Ministry of Science and Innovation and the European Regional Development's funds (FEDER). MSR is co-financed by Consejería de Educación, Juventud y Deporte from Comunidad de Madrid and European Social Fund.

Conflicts of interests: F Abad-Santos and D Ochoa have been a consultant or investigator in clinical trials sponsored by the following pharmaceutical companies: Abbott, Alter, Chemo, Farmalíder, Ferrer, Galenicum, GlaxoSmithKline, Gilead, Janssen-Cilag, Kern, Normon, Novartis, Servier, Teva and Zambon. E Daudén and M. Llamas Velasco have potential conflicts of interest (advisory board member, consultant, grants, research support, participation in clinical trials, honoraria for speaking, and research support). The authors have no other relevant affiliations or financial

involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

POS047 - M

PattRec: An easy-to-use CNV detection tool optimized for targeted NGS panels

Iria Roca[1]; Ana Fernández-Marmiesse[1]; Lorena González-Castro[2]; Sofía Gouveia[1]; Helena Fernández[2]; María Luz Couce[1].

1: Unidad de Diagnóstico y Tratamiento de Enfermedades Metabólicas Congénitas, Hospital Clínico Universitario de Santiago de Compostela, Spain.
2: Gradiant, Centro Tecnológico de Telecomunicaciones de Galicia, Vigo, Spain.

Purpose: Genetic laboratories use custom-commercial targeted next-generation sequencing (tg-NGS) gene panels to identify disease-causing variants. Although the high coverage offered by these tests enable the detection of copy number variants (CNVs), which account for a large proportion of the genetic burden in rare diseases, a robust, easy-to-use tool for automatic CNV detection is still lacking. We describe PattRec, a novel CNV detection tool optimized for tg-NGS data. We also present a novel method for simulating tg-NGS data.

Methodology: PattRec was developed using comprehensive statistical analyses of a broad real tg-NGS dataset (300 DNA samples analyzed using three different tg-NGS panels). In-solution hybridization capture (SureSelect XT; Agilent Technologies) was used as the enrichment method and captured fragments were sequenced as paired-end 100-base reads in the MiSeq platform (Illumina). The performance of the tool was evaluated using a wide range of simulated and real data. The comparison with other previously published packages is also shown. All calculations were performed in the R computing environment. The key features of PattRec are 1) a pre-analytical parameter to identify the most suitable controls for a specific test; 2) a pre-analytical filter to discard polymorphic CNVs; 3) absence of need for same-sex controls, thereby increasing the number of potential controls and decreasing the risk of false positives; 4) SNP filtering to reduce false-positive deletions; 5) creation of in-house CNV MySQL database the first time the program is run, with feature for adding results from subsequent comparisons; 6) easy-to-use, intuitive interface (unlike other methods where the user needs to manage multiple packages in R or Python, in PattRec the user simply has to introduce the bam files for the test and controls, the bedfile containing the regions of interest, the fasta file containing the genome sequence, and the relevant parameters); 7) output in xlsx format (not in plain text like others) with a colour code in addition to p-values for certain parameters to prioritize those CNVs more likely to be pathogenic.

Results: Pattrec is a powerful tool for detecting deletions. With the exception of exomeDepth, PattRec has significantly higher sensitivity than other tools for both simulated and real data. Its performance is comparable to that of exomeDepth for simulated data, but in addition, it offers several novel features not available in this package, such as sex independence, SNP filtering, filtering of common CNVs, an in-house CNV population database, an easy-to-use interface, and an output file that helps user to prioritize potentially real CNVs.

Conclusion: PattRec is significantly more sensitive than other detection tools, especially for small deletions. It incorporates several interesting features, which helps users to prioritize CNVs that are more likely to be real. PattRec is implemented in Java and R, and is packaged as a desktop application with an easy-to-use interface for Ubuntu. A limited-time full trial is freely available at https://goo.gl/forms/uFmrGvFsSQ9MMPIz1.

POS048 - T

Determining the burden of copy number variation in patients with epilepsy

T. K. de Araujo[1], F. R. Torres[1], R. Secolin[1], M. K. M. Alvim[2], C. S. Rocha[1], M. E. Morita[2], C.L. Yasuda[2], B. S. Carvalho[3], F. Cendes[2], I. Lopes-Cendes[1]

[1]Department of Medical Genetics, [2]Department of Neurology, School of Medical Sciences; [3]Department of Statistics, Institute of Mathematics, Statistics and Scientific Computing; University of Campinas (UNICAMP); and the Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, SP, Brazil.

PURPOSE: Mesial temporal lobe epilepsy (MTLE) and genetic generalized epilepsy (GGE) are the most common epilepsies syndromes. Recently, genomic copy number variations (CNVs) have been identified as a risk factor for some types of epilepsy syndromes. The objective of this study is : We aim: **i**) to investigate the distribution of CNVs in patients with MTLE and GGE, as well as in control subjects; **ii**) to evaluate recurrence of CNVs in patients; and **iii**) to identify genes potentially involved in the genetic predisposition to MTLE and GGE within the regions of CNVs found exclusive in patients.

METHODOLOGY: To date, we have studied 750 individuals (340 patients with MTLE, 70 patients with GGE and 340 control subjects). CNVs were assessed by the Genome-Wide Human SNP 6.0 array (Affymetrix, Santa Clara, CA, USA). We analyzed CNVs that were ≥ 100kb and that span at least 25 probes for deletions and 50 probes for duplications.

RESULTS: The analysis of MTLE patients identified 2,246 CNVs. Among all CNVs identified, 20.3% are found only in patients with MTLE and 652 RefSeq genes are affected by these CNVs. Using the Metacore software we found an enrichment of genes associated with neurogenesis development and synaptogenesis, learning or memory, cognition, pre-pulse inhibition, ion transport, protein complex assembly involved in synapse maturation, calcium ion transmembrane transport and single-organism behavior. In patients with GGE, we identified 340 CNVs. Among all CNVs identified, 17.4% are found only in patients with GGE and 59 RefSeq genes are affected by these CNVs. Analysis using Metacore software revealed an enrichment of genes associated with lithium effect on synaptic transmission and autophagy, glutamic acid regulation of dopamine D1A receptor signaling, mitochondrial dysfunction in neurodegenerative diseases, dopamine D2 receptor transactivation of PDGFR in CNS, GABA-B receptor-mediated regulation of glutamate signaling in Purkinje cells, nicotine signaling in dopaminergic neurons, dopamine D2 receptor signaling in CNS. Finally, analysis of the control group identified 2,100 CNVs. Chromosomal regions more affected by CNVs in controls subjects are: 1q21.1, 1q21.2, 1p36.13, 2q11.2, 3q26.1, 4q13.2, 8p11.22, 8p23.1, 10q11.22, 14q11.2, 15q11.2, 15q11.1, 16p11.2, 17q21.31.

CONCLUSION: Our results clearly show that there is an increased burden of CNVs in specific chromosomal regions of the genome in patients with epilepsy. These regions are distinct in patients with MTLE as compared to patients with GGE, indicating that the genetic burden, in these two different epilepsy syndromes, is distinct. We identified structural variants that affect neurodevelopmental genes have a strong impact in crucial neural pathways, leading to epilepsy syndromes. Interestingly, CNVs previously associated with epilepsy syndromes in the literature (1q21.1, 15q11.2, 16p11.2) were also identified in subjects of our control group the significance of which is still unclear.
Support: FAPESP, CNPq.

Exhibition Directory

Exhibit Room A: Lobby

- 10x Genomics
- Agena Bioscience
- Interactive Biosoftware
- LGC
- PacBio
- Sophia Genetics
- Qiagen

Exhibit Room B: Quintana Room, Lobby level

- Illumina
- Phenosystems
- Saphetor
- Reference Laboratory Genetics

Poster Directory

- Posters P001 - P018: Quintana room, Lobby level
- Posters P019 - P048: Lobby

Sponsors & Exhibitors

10x Genomics meets the critical need for long range, structural and cellular information, with an innovative system that transforms short-read sequencing technologies. Our Chromium™ System supports comprehensive genomics and high-throughput single cell transcriptomics. It enables researchers to discover previously inaccessible genomic information at unprecedented scale, including phased structural variants, phased single nucleotide variants, and dynamic gene expression of individual cells—while leveraging their existing sequencing systems and workflows.

www.10xgenomics.com

Agena Bioscience is a San Diego, CA based life sciences and clinical diagnostics company that recently acquired the Bioscience business of Sequenom, Inc. and is now offering the MassARRAY® System. The system is a highly sensitive, quantitative method for nucleic acid detection via mass spectrometry for high-throughput genotyping and mutation profiling for cancer and other disease research, companion diagnostics, pharmacogenomics, molecular blood group typing, epigenetics, clinical genetics, agrigenomics, and molecular sample identification for bio-banking.

http://agenabioscience.com

AstraZeneca is a research based pharmaceutical company. In the area of oncology AstraZeneca is working to redefine the cancer treatment paradigm by delivering life-changing medicines to help address unmet clinical needs in a host of cancers and one day help eliminate the disease as a cause of death.

www.astrazeneca.com

The Irys platform from **BioNano Genomics** provides unprecedented insights into whole-genome biology. Genome maps are generated from massively parallel single-molecule visualization of extremely long DNA. These maps provide dense genome-wide anchor points for ordering and orienting sequencing contigs or scaffolds to greatly increase completion and accuracy of de novo assemblies. Structural variants and repeats are measured directly within long “reads” for comprehensive analysis of what has been referred to as the inaccessible genome. For more see www.bionanogenomics.com

www.bionanogenomics.com

At **Illumina**, our goal is to apply innovative technologies and revolutionary assays to the analysis of genetic variation and function, making studies possible that were not even imaginable just a few years ago. These studies will help make the realization of personalized medicine possible. With such rapid advances in technology taking place, it is mission critical to have solutions that are not only innovative, but flexible, scalable, and complete with industry-leading support and service. As a global company that places high value on collaborative interactions, rapid delivery of solutions, and prioritizing the needs of its customers, we strive to meet this challenge. Illumina's innovative, Next Generation Sequencing as well as array-based solutions for genomic analysis serve as tools for disease research, drug development, and the development of molecular tests in the clinic.

www.illumina.com

Interactive Biosoftware is the creator of the Alamut® Software Suite offering a comprehensive solution to the complex tasks of genomic variants annotation, filtration, interpretation and reporting.

Interactive Biosoftware is changing genetic diagnostics and research as we know it by simplifying the mutation interpretation process, while saving scientists’ time, improving outcome quality and enhancing productivity.

www.interactive-biosoftware.com

LGC is a global leader in delivering genomic solutions for research, diagnostics, and applied markets. We provide best-in-class products supporting quantitative and end-point PCR (KASP and BHQ® probes) and RNA fluorescence in situ hybridization (Stellaris® RNA FISH probes). We offer state-of-the-art instrumentation for extraction (Oktopure™) and a fully integrated PCR platform (IntelliQube®). Our innovative technologies also power lab services for genotyping, DNA extraction, and sequencing are also available.

www.lgcgroup.com/genomics

PacBio® Single Molecule, Real-Time (SMRT®) Sequencing provides full access to human genomic variation through unmatched read length, uniform coverage and exceptional accuracy. With average reads lengths greater than 10 kb, PacBio DNA sequencing data reveals previously hidden structural variants and produces direct variant phasing information across haplotype blocks, in both whole genome and targeted applications. With SMRT Sequencing Systems, scientists gain new insight into the genetic basis of disease heritability and a more comprehensive view of human genetic variation.

www.pacb.com

Phenosystems develops user-friendly software since 2002. Our software suite for molecular diagnostics comprises: GensearchNGS dedicated to mutation detection and interpretation on data from gene panels, Whole Exomes and Whole Genome; Gensearch for capillary DNA sequencing, GensearchGT Genotyping, GensearchHIV. Major features are high specificity and sensitivity, advanced tools to support variant interpretation (frameshifts, splice prediction, connection to LSDBs and to CaféVariome.org). All have been developed together with leading diagnostics laboratories in Europe, packing powerful tools in an extremely user friendly interface.
Contact email: contact@phenosystems.com

www.phenosystems.com

QIAGEN is the leading global provider of Sample to Insight solutions to transform biological materials into valuable molecular insights. QIAGEN sample technologies isolate and process DNA, RNA and proteins from blood, tissue and other materials. Assay technologies make these biomolecules visible and ready for analysis. Bioinformatics software and knowledge bases interpret data to report relevant, actionable insights. Automation solutions tie these together in seamless and cost-effective molecular testing workflows.

www.qiagen.com

Reference Laboratory Genetics is the Genetics and Molecular Diagnosis Division of ReferenceLaboratory. It is nowadays the biggest laboratory of Genetics in Europe and guarantees the accomplishment of any genetic study required in his facilities with own resources. The patient is our priority. We consider each case individually and suggest to carry on the most appropriate genetic studies. All the genetic reports we issue are customized and are adapted to EMQN's recommendations.
Our Head of Genetic Counselor is Dr. Esther Geán.

http://reference-laboratory.es

Saphetor, the genome interpreter.
Saphetor provides end-to-end solution analysis for next generation DNA sequencing for research and clinical use. We seamlessly combine the world's leading molecular databases, coupled with advanced software. To empower the research/medical professional to accurately identify the cause of cancer or genetic disease, and connect the patient to treatment options and clinical trials, we provide timely reports based on the most up-to-date information from academic and clinical research tailored to the patient's DNA data.

http://saphetor.com

Sophia Genetics is the Global Leader in Data Driven Medicine. Our team of talented experts in Clinical Genomics, Machine Learning and Bioinformatics built Sophia DDM®, the most advanced analytical platform for clinical diagnostics. Over 175 leading hospitals use Sophia DDM® to diagnose thousands of patients each week in the fastest, most accurate possible way.

www.sophiagenetics.com

Attendee List at 23rd May 2017

A

Dr. Joao Abade
Sophia Genetics
St-Sulpice, Switzerland
jabade@sophiagenetics.com

Dr. Monica Albarca Aguilera
Saphetor
Customer Operations
Lausanne, Switzerland
monica.albarca@saphetor.com

Miss Vanessa S Almeida
University of Campinas - Unicamp
Department of Medical Genetics/
Faculty of Medical Science
Campinas, Brazil
va.salmeida25@gmail.com

Dr. Isabel Alonso
IBMC, i3S
UnIGENe
Porto, Portugal
ialonso@ibmc.up.pt

Miss Aitana Alonso González
Fundación Galega de Medicina
Xenomica
Genética
Santiago de Compostela, Spain
aitana.alonso.glez@gmail.com

Mr. Faisal S Alquaddoomi
UCLA
Computer Science
Los Angeles, USA
falquaddoomi@gmail.com

Dr. Jorge Amigo
Universidad de Santiago de
Compostela
Grupo de Medicina Xenómica
Santiago de Compostela, Spain
jorge.amigo@usc.es

Mrs. Juan Ansede
CEGEN-PRB2-ISCIII-USC
CEGEN
Santiago de Compostela, Spain
juan.ansede@usc.es

Dr. Swaroop Aradhya
InVitae
Medical Genetics
San Francisco, USA
swaroop.aradhya@invitae.com

Dr. David Atlan
Phenosystems SA
Research
Braine le Chateau, Belgium
phenosystems@gmail.com

Dr. Ana Aza
Labclinics SA
Marketing
Barcelona, Spain
marketing@labclinics.com

B

Sofia Isabel Barbosa Sousa Gouveia
Barcelona, Spain
jcebrian@agora-events.com

Dr. Francisco Barros
FPGMX
Molecular Medicine
Santiago de Compostela, Spain
francisco.barros@usc.es

Mr. Stephan Bauer
LGC
Genomics
Hoddesdon, UK
stephan.bauer@lgcgroup.com

José Maria Belloso Sanchez
Sophia Genetics
St-Sulpice, Switzerland
jmbelloso@sophiagenetics.com

Dr. Anna Benet-Pages
MGZ - Medizinisch Genetisches
Zentrum
NGS
Munich, Germany
benet-pages@mgz-muenchen.de

Mr. Martin Berg
Cape Canaveral, USA
bioengineer@comcast.net

Ms. María Lara Besada Cerecedo
Health Research Institute of Santiago de
Compostela
Group of Genetics and Developmental
Biology of Renal Diseases
Santiago de Compostela, Spain
larbescer@hotmail.com

Ms. Patricia Blanco
Hospital Clínico de Santiago
Grupo de Neurogenética IDIS
Santiago de Compostela, Spain
pba233@gmail.com

Mrs. Ana Blanco Perez
Fundacion Publica Galega de Medicina
Xenomica
Oncological department
Santiago de Compostela, Spain
ana.blanco@usc.es

Alejandro José Blanco Verea
Instituto de Investigación Sanitaria
Santiago de Compostela, Spain
alejandroj.blanco@usc.es

Dr. André Blavier
Interactive Biosoftware
Rouen, France
ablavier@interactive-biosoftware.com

Sven Bocklandt
Bionano Genomics
San Diego, USA
sbocklandt@bionanogenomics.com

Ms Jana Bohmova
University Hospital
Medical Genetics and Fetal Medicine
Olomouc, Czech Republic
bohmovajana@seznam.cz

Dr. María Brión Martínez
Instituto de Investigación Sanitaria
Xenética de enfermidades
cardiovasculares e oftalmolóxicas
Santiago de Compostela, Spain
maria.brion@usc.es

Ms. Elena Buena Atienza
Institute for Ophthalmic Research
Molecular Genetics Lab
Tuebingen, Germany
elenabuena3@gmail.com

Dr. Henk Buermans
Leiden University Medical Center
Human Genetics / LGTC
Leiden, The Netherlands
h.buermans@lumc.nl

Ms. Anna Bukaeva
Petrovsky Russian Research Centre of
Surgery
Medical Genetics Laboratory
Moscow, Russia
annbukaeva@gmail.com

Prof. Sir John Burn
Newcastle University
Institute of Genetic Medicine
Newcastle upon Tyne, UK
john.burn@newcastle.ac.uk

Dr. Ruth Burton
Qiagen
Advanced Genomics
Manchester, UK
ruth.burton@qiagen.com

C

Mrs. Pilar Cacheiro
University of Santiago de Compostela
Grupo de Mecicina Xenómica/
Grupo Neurogenética IDIS
Santiago de Compostela, Spain
pilar.cacheiro@usc.es

Pilar Camaño
University of Santiago de Compostela
Santiago de Compostela, Spain
PILAR.CAMANOGONZALEZ@osakidetza.eus

Prof. Angel Carracedo
University of Santiago de Compostela
Santiago de Compostela, Spain
angel.carracedo@usc.es

Dr. Paola Carrera
IRCCS San Raffaele Scientific Institute
Dept of Genetics and Cell Biology
Milano, Italy
carrera.paola@hsr.it

Miss Cristina Castro Fernández
Fundación Pública Galega de Medicina Xenómica
Neurogenetics group
Santiago de Compostela, Spain
cristinacastrofdez@gmail.com

Mr. Manuel Cendagorta-Galarza Lopez
Instituto Tecnológico y de Energías Renovables
Gerencia
Granadilla de Abona, Spain
cendagorta@iter.es

Dr. Stephen J. Chanock
National Cancer Institute
Division of Cancer Epidemiology and Genetics
Bethesda, USA
chanocks@mail.nih.gov

Dr. George Chong
Jewish General Hospital
Molecular Pathology
Montreal, Canada
george.chong@mcgill.ca

Mr. Omar F Cruz-Correa
Universidad Nacional Autónoma de México
Doctorado en Ciencias Bioquímicas
Mexico City, México
omar.cruzcorrea@gmail.com

D

Dr. Runa Daniel
Victoria Police Forensic Services Department
Office of the Chief Forensic Scientist
Melbourne, Australia
runa.daniel@police.vic.gov.au

Prof. Johan T. den Dunnen
Leiden University Medical Center
Leiden, Netherlands
ddunnen@humgen.nl

Laurence Desmyter
HOPITAL ERASME
Brussels, Belgium
laurence.desmyter@erasme.ulb.ac.be

Miss Ana Diaz de Usera
Instituto Tecnológico y de Energías Renovables
Genomics' Division
Granadilla de Abona, Spain
anadidu07@gmail.com

Mrs. Mev Dominguez-Valentin
Oslo University Hospital
Tumor Biology
Oslo, Norway
mev.dominguez.valentin@rr-research.no

Prof. Peter J Donnelly
University of Oxford
Wellcome Trust Centre for Human Genetics
Oxford, UK
donnelly@well.ox.ac.uk

Dr. Joaquin Dopazo
Fundacion Progreso y Salud
Clinical Bioinformatics Research Area
Sevilla, Spain
joaquin.dopazo@juntadeandalucia.es

Zelie Dubreucq
Sophia Genetics
St-Sulpice, Switzerland
zdubreucq@sophiagenetics.com

Prof. Bernd Dworniczak
University hospital münster
Human genetics
Münster, Germany
dwornic@uni-muenster.de

E

Dr. Rosemary Ekong
UCL (University College London)
Genetic, Evolution & Environment
London, UK
r.ekong@ucl.ac.uk

Dr. Per O Ekstrom
Institute for Cancer Research
The Norwegian Radium hospital
Tumor biology
Oslo, Norway
pok@rr-research.no

Mr. Kevin W Ericksen
Monash Health
Pathology
Melbourne, Australia
k.ericksen@monashhealth.org

Miss Uxia Saraiva Uxía Esperón Moldes
Fundación Pública Galega de Medicina Xenómica
Medicina Legal
santiago de compostela, Spain
uxiasaraiba@gmail.com

Ms. Eva Maria Esteban
FPGMX
Santiago de Compostela, Spain
Eva.Maria.Esteban.Cardenosa@sergas.es

F

Dr. Aída B. Falcón de Vargas
Hospital Vargas de Caracas
Clinical Genetics Unit
Caracas, Venezuela
aidafalvar@gmail.com

Ana Fernandez Marmiesse
Barcelona, Spain
JCEBRIAN@AGORA-EVENTS.COM

Dr. Ceres Fernández Rozadilla
Fundación Ramón Domínguez
FPGMX
Santiago de Compostela, Spain
ceres.fernandez.rozadilla@gmail.com

Dr. Juan Fernandez Tajes
Wellcome Trust Centre for Human Genetics
Nuffield Department of Medicine
Oxford, United Kingdom
jfertaj@well.ox.ac.uk

Dr. Sandra E Filippini
Universidad de Santiago de Compostela
FPGMX
Santiago de Compostela, Spain
sandraefilip@gmail.com

Miss Aisling C Fiuza
Universidad de Santiago de Compostela
Departamento de Ciencias Foresees Anatomía Patolóxica, Xinecoloxía e Obstetricia e Pediatría
Santiago de Compostela, Spain
aislingfiuza@gmail.com

Prof. Carlos Alberto A Flores Infante
Instituto Técnológico y de Energías Renovables
Genómica
Granadilla de Abona, Spain
cflores@ull.edu.es

Miss Cristina Fortuno
QIMR Berghofer
Genetics & Computational Biology Department
Brisbane city, Australia
cristina.fm@hotmail.es

G

Dr. Maria Garcia-Murias
IDIS
Neurogenetics
Santiago de Compostela, Spain
mariagmurias@gmail.com

Dr. Mette Gaustadnes
Aarhus University Hospital
Department of Molecular Medicine
Aarhus N, Denmark
mette.gaustadnes@clin.au.dk

Mª Esther Geán
Reference Laboratory Genetics
Barcelona, Spain
egean@referencelaboratory.es

Prof. Sujoy Ghosh
Duke-NUS Medical School
Centre for Computational Biology
Singapore
sujoy.ghosh@duke-nus.edu.sg

Mr. Steve Glavas
10x Genomics
Sales
Stockholm, Sweden
steve.glavas@10xgenomics.com

Miss Rafaela Gonzalez Monterlongo
Instituto Técnológico y de Energías Renovables
Genómica
Granadilla de Abona, Spain
rfgmont@gmail.com

Dr. Javier González Peñas
Universidad de Santiago de Compostela
Molecular medicine
Santiago de Compostela, Spain
javipenhas@gmail.com

Dr. Ana Gorostidi
Biodonostia
Genomic Platform
Donostia, Spain
ana.gorostidi@biodonostia.org

Dr. Ivo G Gut
Centro Nacional de Análisis Genómico, CNAG-CRG
Management
Barcelona, Spain
ivo.gut@cnag.crg.eu

H

Dr. Jennifer Harrow
Illumina
Population Sequencing
Cambridge, UK
jharrow@illumina.com

Dr. Reece Hart
Invitae
San Francisco, USA
reecehart@gmail.com

Mr. Luke E Hickey
PacBio
Marketing
Menlo Park, USA
lhickey@pacb.com

Annika Hjerdin Panagopoulos
Oslo University Hospital
Oslo, Norway
annpan@ous-hf.no

Dr. Georgina Hollway
Genome.One
Clinical Interpretation
Darlinghurst, Australia
georgina.hollway@genome.one

Ms. Rania Horaitis
Meeting Makers
Event Manager
Melbourne, Australia
rania@meeting-makers.com

Aleksandra Hoven
Oslo University Hospital
Oslo, Norway
asilye@ous-hf.no

K

Dr. Jitka Kadlecová
Cytogenetic Laboratory Brno
Molecular diagnostic laboratory
Brno, Czech Republic
jkd@centrum.cz

Miss Tânia K Kawasaki de Araujo
University of Campinas
Medical Genetis - Faculty of Medical Science
Campinas, Brazil
taniakawasaki@gmail.com

Ms. Anabel Kearney
South West Sydney Pathology Liverpool Haematology
Sydney, Australia
anabel.kearney@sswahs.nsw.gov.au

Ms. Triin Kikas
University of Tartu
Institute of Molecular and Cell Biology
Tartu, Estonia
triinc7@gmail.com

Dr. Wioletta Krysa
Institute of Psychiatry and Neurology
Department of Genetics
Warsaw, Poland
krysa@ipin.edu.pl

Dr. Erika Küchler
University of Sao Paulo
Pediartic Dentistry
Ribeirao Preto, Brazil
erikacalvano@gmail.com

Mr. John Kuijpers
PacBio
Sales
Tilburg, The Netherlands
jkuijpers@pacb.com

L

Prof. Georgia Lahr
University Medical Center Ulm
Molecular Diagnostics Laboratory
Diagnostic Laboratories Department of Pediatrics and Adolescent Medicine
Ulm, Germany
georgia.lahr@uni-ulm.de

Mrs. Séverine Lair
Interactive Biosoftware
Bioinformatics
Rouen, France
slair@interactive-biosoftware.com

Dr. Ryan Lamont
Alberta Children's Hospital
Molecular Diagnostics Lab
Calgary, Canada
ryan.lamont@albertahealthservices.ca

Prof. Doron Lancet
Weizmann Institute
Molecular Genetics
Rehovot, Israel
doron.lancet@weizmann.ac.il

Dr. Andreas Laner
MGZ - Medical Genetics Centre
Molecular Genetics
Munich, Germany
laner@mgz-muenchen.de

Mrs. Ana Latorre Pellicer
Universidade de Santiago de Compostela
Medicina Xenomica
Santiago de Compostela, Spain
ana.latorre@usc.es

Mr. Diego H Levi
Ministry of Science and Technology
Precisión Medicine
Belgrano, Argentina
dlevi99@yahoo.com

Dr. Melanie Locher
MGZ - Medical Genetics Center
Molecular Genetics
Munich, Germany
melanie.locher@mgz-muenchen.de

Dr. Lourdes Loidi
Fundación Pública galega de Medicina Xenómia
Molecular Medicine
Santiago de Compostela, Spain
lourdes.loidi.fernandez@sergas.es

Prof. Mabel Loza García
University of Santiago de Compostela
Pharmacology
Santiago de Compostela, Spain
mabel.loza@usc.es

M

Mr. Jose Miguel Moreno Salazar
Institute of Technology and Renewable Energy (ITER)
Division of Genomics
Granadilla de Abona, Spain
jlorsal@gmail.com

Mrs. Olalla Maroñas
Universidad de Santiago de Compostela
Medicina Xenómica
Santiago de Compostela, Spain
olalla.maronas@hotmail.com

Dr. Elisabeth Maurer
Medical University Innsbruck
Division of Human Genetics
Innsbruck, Austria
elisabeth.maurer@i-med.ac.at

Dr. Andrea Michenkova
Center of Prenatal Diagnostics
Center of prenatal diagnostics
Brno, Czech Republic
andreamichenkova@seznam.cz

Dr. Jusaku Minari
Osaka University
Department of Biomedical Ethics and Public Policy
Osaka, Japan
minari@eth.med.osaka-u.ac.jp

Dr. Yury Monczak
McGill University Health Center
Pathology
Montreal, Canada
mikroba1@yahoo.ca

Ms. Ciara M Mulhern
UCL Great Ormond Street Institute of Child Health
Infection inflammation and rheumatology
London, UK
ciara.mulhern.16@ucl.ac.uk

N

Dr. Muhammad Israr I Nasir
Liaquat National Hospital and Medical College
Molecular Pathology
Karachi, Pakistan
israr.nasir@lnh.edu.pk

Birgit Neitzel
MGZ - Medical Genetics Center
Munich, Germany
Birgit.Neitzel@mgz-muenchen.de

Dr. ChongHan Ng
Multimedia University
Faculty of Information and Science Technology
Bukit Beruang, Malaysia
chonghan.ng@gmail.com

Dr. Andrew J Nightingale
EMBL-EBI
Bioinformatician
Hinxton, UK
anight@ebi.ac.uk

O

Dr. Noriko Ohashi
Osaka University Graduate School of Medicine
Department of Biomedical Ethics and Public Policy
Suita, Japan
noriko@eth.med.osaka-u.ac.jp

Prof. Daniela Oliveira
Federal University of Alfenas
Clinic and Surgery
Alfenas, Brazil
barrosodaniela@hotmail.com

Ms. Soledad Otero
Universidade de Santiago de Compostela
Medicina Xenómica, Universidade de Santiago de Compostela
Santiago de Compostela, Spain
msoledad.otero@usc.es

Dr. María C C Ovejero-Benito
Instituto de Investigación del Hospital Universitario de La Princesa
Clinical Pharmacology
Madrid, Spain
covejero@salud.madrid.org

P

Mr. Joris Parmentier
LGC
Genomics
Hoddesdon, UK
joris.parmentier@lgcgenomics.com

Prof. Walther Parson
Medical University of Innsbruck
Institute of Legal Medicine
Innsbruck, Austria
walther.parson@gmail.com

Pablo Pavon
Reference Laboratory Genetics
Barcelona, Spain
ppavon@referencelaboratory.es

Dr. Christopher Phillips
University of Santiago de Compostela
Forensic Genetics Unit, Institute of Forensic Sciences
Santiago de Compostela, Spain
c.phillips@mac.com

Dr. Ekaterina Pomerantseva
Center of Genetics and Reproductive Medicine Genetico LLC
Genetic laboratory
Moscow, Russia
e.pomerantseva@gmail.com

Mr. Rutger Prins
Bionano Genomics
Sales
San Diego, USA
rprins@bionanogenomics.com

Q

Ms. Dora Quest
LGC
Genomics
Hoddesdon, UK
dora.quest@lgcgroup.com

Dr. Beatriz Quintáns
Instituto de Investigación Sanitaria de Santiago de Compostela
Neurogenetics group
Santiago de Compostela, Spain
beaquintans@gmail.com

Miss Rita Quintas Rey
Fundacion Publica Galega de Medicina Xenomica
Santiago de Compostela, Spain
rita.quintas.rey@gmail.com

Dr. Ines Quintela
CEGEN
Santiago de Compostela, Spain
ines.quintela@usc.es

R

Dr. Gunaretnam Rajagopal
Janssen Research & Development
Discovery Sciences
Spring House, USA
grajagop@its.jnj.com

Eva Ramos Luis
Instituto de Investigación Sanitaria
Santiago de Compostela, Spain
evaramosluis@gmail.com

Dr. Pablo Rana
Fundacion Publica Galega De Medicina Xenomica
Hereditary Breast And Ovarian Cancer
Santiago De Compostela, Spain
paulusranae@gmail.com

Prof. Gunnar Ratsch
ETH Zurich
Computer Science
Zurich, Switzerland
raetsch@inf.ethz.ch

Mr. Paulo Refinetti
Ecole Polytechnique Federale Lausanne
Applied Statistics
Lausanne, Switzerland
paulo.refinetti@epfl.ch

Prof. Peter Robinson
The Jackson Laboratory for Genomic Medicine
Computational Biology
Farmington, USA
peter.robinson@jax.org

Iria Rock Otero
Barcelona, Spain
jcebrian@agora-events.com

Dr. Julio Rodríguez
Instituto de Investigación Sanitaria
Universidade de Santiago de Compostela
Medicina Xenómica
Santiago de Compostela, Spain
julrodr80@gmail.com

Dr. Marta Rodríguez-Balada
Unitat de Consell Genètic (Oncologia)
Institut d'Oncologia de la Catalunya Sud
Hospital Universitari Sant Joan de Reus
Reus, Spain
mrodriguez@grupsagessa.com

Prof. Peter K Rogan
University of Western Ontario
Biochemistry
London, Canada
progan@uwo.ca

Ms. Vibeke Wethe Rognlien
Oslo University Hospital
Department of Medical Genetics
Oslo, Norway
vibwet@ous-hf.no

Dr. Anna M Rohlin
Sahlgrenska University Hospital
Clinical Genetics
Göteborg, Sweden
anna.rohlin@vgregion.se

Dr. Barbara Roig Bourgine
Hosp. Univ. Sant Joan de Reus
Oncology
Reus, Spain
Barbara.roig@grupsagessa.com

Dr. Anne Ronan
HNELHD
Hunter Genetics Unit
Newcastle NSW, Australia
anne.ronan@hnehealth.nsw.gov.au

Mrs. Irina Royo
Reference Laboratory
Molecular Biology
Hospitalet de llobregat
Barcelona, Spain
iroyo@referencelaboratory.es

Dr. Clara Ruiz-Ponte
Fundacion Publica Galega Medicina Xenomica
Medicina Molecular
Samtiago de Compostela, Spain
clara.ruiz.ponte@usc.es

S

Ms. Mari Sæther
St. Olavs Hospital
Medical genetics
Trondheim, Norway
Mari.Sether@stolav.no

Miss Marta Santamariña
CIBERER
FPGMX
Santiago de Compostela, Spain
santamarinapena@gmail.com

Dr. Alexander F Sartori
Agena Bioscience
Applications Development
Hamburg, Germany
alexander.sartori@agenabio.com

Dr. Christa M. Schmidt
St. Olav hospital
Medical Genetics
Trondheim, Norway
christa.schmidt@stolav.no

Mr. Stefan Schrader
Agena Bioscience
Sales
Hamburg, Germany
stefan.schrader@agenabio.com

Ms. Daniela Seminara
National Cancer Institute
National Cancer Institute
Rockville, USA
seminard@mail.nih.gov

Mrs. Anna Shestak
Petrovsky Russian Research Centre of Surgery
Medical Genetics Laboratory
Moscow, Russia
anna.shestak87@gmail.com

Dr. Robert A Smith
Institute of Health and Biomedical Innovation
Queensland University of Technology
Genomics Research Centre
School of Biomedical Sciences
Kelvin Grove, Australia
r157.smith@qut.edu.au

Dr. María Jesús Sobrido
Fundación Publica Galega de Medicina Xenómica
Neurogenetics
Santiago de Compostela, Spain
ssobrido@telefonica.net

Dr. Beatriz Sobrino
Fundación Publica Galega de Medicina Xenómica
Xenomica
Santiago de Compostela, Spain
beatriz.sobrino@usc.es

Dr. Angela R Solano
CEMIC and INBIOMED UBA/CONICET
Biochemistry
CABA, Argentina
drsolanoangela@gmail.com

Dr. Susana M Sousa
IBMC
CGPP
Porto, Portugal
susana.mendes@ibmc.up.pt

Prof. Thierry Soussi
INSERM UMRS 1138
Equipe G. KROEMER
Paris, France
lynda.bennaci@upmc.fr

Dr. Daniela Steinberger
bio.logis Center for Humangenetics
Frankfurt, Germany
daniela.steinberger@bio.logis.de

Dr. Alessandro Stella
Università degli Studi di Bari Aldo Moro
Scienze Biomediche e Oncologia Umana
Bari, Italy
alessandro.stella@uniba.it

Dr. Dominik Strapagiel
University of Lodz
Department of Molecular Biophysics
Lodz, Poland
dominik.strapagiel@biol.uni.lodz.pl

Mr. José Javier Suárez Rama
Universidade de Santiago de Compostela
CIMUS
Santiago de Compostela, Spain
josejavier.suarez@usc.es

Dr. Anna Sulek
Institute of Psychiatry and Neurology
Department of Genetics
Warsaw, Poland
suleka@ipin.edu.pl

Prof. Moon Woo Sung
Seoul National University Hospital
Laboratory Medicine
Seoul, South Korea
mwseong@snu.ac.kr

Prof. Jordi Surralles
Hospital Sant Pau/
Universitat Autònoma de Barcelona
Genetics
Barcelona, Spain
jordi.surralles@uab.es

Mrs. Dianne E Sylvester
Kids Research Institute
Children's Cancer Research Unit
Westmead, Australia
dianne.sylvester@health.nsw.gov.au

T,V

Dr. Clare A Turnbull
Genomics England
London, UK
clare.turnbull@icr.ac.uk

Ms. Vivien Vasic
Hudson Institute of Medical Research
Core Facilities
Clayton, Australia
vivien.vasic@hudson.org.au

Dr. Ana Vega
Hospital CI
FPGMX
Santiago de Compostela, Spain
ana.vega@usc.es

Dr. Maria Carmen Vidal Lampurdanés
Hospital Son Espases
Sequencing Unit
Palma de Mallorca, Spain
mariac.vidal@ssib.es

Dr. Catheline N Vilain
ULB Center of Human Genetics
Brussels, Belgium
cavilain@ulb.ac.be

Dr. Radek Vodicka
Palacky University
Medical Genetic
Olomouc, Czech Republic
vodickar@fnol.cz

Dr. Dita Vrbicka
U.S.G.POL s.r. o
Laboratory of molecular genetics
Olomouc, Czech Republic
dita.vrbica@email.cz

W

Mrs. Françoise Wilkin
Hospital Erasme
Genetics
Brussels, Belgium
francoise.wilkin@erasme.ulb.ac.be

Tobias Wohlfrom
MGZ - Medical Genetics Center
Munich, Germany
tobias.wohlfrom@mgz-muenchen.de

Mr. Beat Wolf
HES-SO
Informatics
Fribourg, Switzerland
beat.wolf@hefr.ch

Y

Dr. Natsuko Yamamoto
Osaka University
Department of Biomedical Ethics and
Public Policy
Graduate School of Medicine
Suita, Japan
nyamamot@eth.med.osaka-u.ac.jp

Dr. Fengtang Yang
Wellcome Trust Sanger Institute
Cytogenetics
Cambridge, United Kingdom
fy1@sanger.ac.uk