EUROPEAN SOCIETY OF HUMAN GENETICS

European School of Genetic Medicine

3rd Course in

Next Generation Sequencing

Bertinoro, Italy, May 7-10, 2014

Bertinoro University Residential Centre
Via Frangipane, 6 – Bertinoro

Course Director:
J. Veltman (Nijmegen, The Netherlands)

Course Co-Directors:
Gilissen C. (Nijmegen, The Netherlands), Pippucci T. (Bologna, Italy)

EUROPEAN SCHOOL OF GENETIC MEDICINE

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

CENTRO RESIDENZIALE UNIVERSITARIO Bertinoro

FONDAZIONE INTERNAZIONALE MENARINI

3rd Course in

Next Generation Sequencing

Bertinoro, Italy, May 7-10, 2014

CONTENTS

3rd COURSE IN NEXT GENERATION SEQUENCING

Bertinoro University Residential Centre
Bertinoro (Italy), May 7-10, 2014

Arrival: Tuesday May 6

Wednesday, May 7

Morning Session: Introduction to Next Generation Sequencing

9.00 – 9.15	Introduction to the course **Giovanni Romeo**
9.15 – 10.15	Next Generation Sequencing basics **Joris Veltman**
10.15 – 10.45	**Coffee Break**
10.45 – 11.45	Bioinformatic basics **Christian Gilissen**
11.45 – 12.45	Bioinformatic strategies & ontology's **Peter Robinson**
12.45– 13.30	**Lunch Break**

Afternoon Session:

13.30 –14.00	Poster Viewing Session
14.00 – 16.00	Concurrent Workshops: Computer practical: NGS Basics (**C. Gilissen** & **T. Pippucci**) Workshops by speakers
16.00-16.30	**Coffee Break**
16.30 – 18.00	Concurrent Workshops Computer practical: NGS Basics (**C. Gilissen** & **T. Pippucci**) Workshops by speakers

Thursday, May 8

Morning Session: Applications to disease gene identification & diagnostics

9.00 – 10.00	Targeted breast cancer diagnostics **Gert Matthijs**
10.00 – 11.00	Exome diagnostics in intellectual disability **Anita Rauch**
11.00 – 11.30	**Coffee Break**
11.30- 12.30	De novo mutations in human genetic disease **Joris Veltman**
12.30 – 13.30	NGS and non invasive prenatal diagnosis **C. Boustred**
13.30 – 14.30	**Lunch Break**

Afternoon Session:

14.30 – 16.00	Concurrent Workshops Computer practical: Disease gene identification (**C. Gilissen** & **T.Pippucci**) Workshops by speakers
16.00-16.30	**Coffee Break**
16.30 – 18.00	Concurrent Workshops Computer practical: Disease gene identification (**C. Gilissen** & **T.Pippucci**) Workshops by speakers

Friday, May 9

Morning Session: Applications in common disease & cancer

9.00 – 10.00	NGS in population genetics and complex diseases **Paul de Bakker**
10.00 – 11.00	Whole genome sequencing of 3,500 individuals in an isolated population: theSardinia Sequencing Project **Carlo Sidore**
11.00 – 11.30	**Coffee Break**
11.30 – 12.30	Cancer genome sequencing **Ian Tomlinson**
12.30 – 13:30	Exome sequencing to study rare and common variation in diabetes **Amélie Bonnefond**
13:30 – 14.30	**Lunch Break**

Afternoon Session:

14.30 –15.00	Poster Viewing Session (or CLCbio/Cartegenia demonstration?)
15.00 – 16.30	Concurrent Workshops Computer practical: Diagnostic NGS (**C. Gilissen** & **T. Pippucci**) Workshops by speakers
16.30-17.00	Coffee Break
17.00 – 18.30	Concurrent Workshops Computer practical: Diagnostic NGS (**C. Gilissen** & **T. Pippucci**) Workshops by speakers

Saturday, May 10

Morning Session: Genome technologies

9.00 – 10.00	Retrieval and annotation of human mitochondrial mutations from NGS experiments. **F. Calabrese**
10.0	
11.0 – 11.00	Analysis of mitochondrial genomes **A. Ameur**
11.00 – 11.30	**Coffee Break**
11.30 – 12.00	Best Posters Presentations by students
12.00 – 12.30	Wrapping up of the course
12.30	**Lunch**

Departure

An application has been made to the EACCME for CME accreditation of this event.
*The application for European accreditation has been granted **21 European CME credits** (ECMEC) by the European Accreditation Council for Continuing Medical Education (EACCME).*

ABSTRACTS OF LECTURES

Friday, May 7

Next Generation Sequencing basics

Joris A.Veltman
Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

There is considerable variation between the genetic code of two individuals, both at the single nucleotide and at the structural level. Identifying and studying the consequences of these variations, a core activity in human genetics research, is driven by technological innovations. Currently we are in the midst of one of the greatest technological revolutions in genomics. Novel DNA sequencing methods are dramatically increasing sequencing throughput to a level where it is soon possible to rapidly sequence an individual genome for an affordable price. If properly established, whole genome sequencing will have a major impact on the entire field of medicine; All genomic variation that can be linked to disease is detectable in a single experiment!
In this presentation I will introduce next generation sequencing technology and discuss its development and advantages over traditional sequencing technologies.

Recommended reading:

1. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 26: 1135-45 (2008).
2. Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. The next-generation sequencing technology and application. Protein Cell 1: 520-36 (2010). Review.
3. Tucker T, Marra M, Friedman JM. Massively parallel sequencing: the next big thing in genetic medicine. Am J Hum Genet 85:142-54 (2009). Review.
4. Ashley EA, Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, Dudley JT, Ormond KE, Pavlovic A, Morgan AA, Pushkarev D, Neff NF, Hudgins L, Gong L, Hodges LM, Berlin DS, Thorn CF, Sangkuhl K, Hebert JM, Woon M, Sagreiya H, Whaley R, Knowles JW, Chou MF, Thakuria JV, Rosenbaum AM, Zaranek AW, Church GM, Greely HT, Quake SR, Altman RB. Clinical assessment incorporating a personal genome. Lancet 375: 1525-35 (2010).

Bioinformatic basics

Christian Gilissen
Department of Human Genetics -
Radboud University Nijmegen Medical Centre, the Netherlands

Next Generation Sequencing (NGS) technologies have revolutionized the field of medical genetics research by generating large numbers of DNA sequences within a matter of days at very low cost. Next generation sequencing is being used extensively to search for Mendelian disease genes in an unbiased manner by sequencing the entire protein-coding sequence, known as the exome, or even the entire human genome.[1] Increasingly, NGS is also being applied for the diagnosis of patients with genetically heterogeneous disorders, where sequencing of all individual disease genes in infeasible.[2,3]

Because of the large amounts of data that are being generated, bioinformatics plays an increasingly important role. In this talk I will focus on the basic bioinformatic concepts, data formats and pitfalls of analyzing NGS data from resequencing experiments for applications in research and diagnostics.[4]

[1] Unlocking Mendelian disease using exome sequencing. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Genome Biol. 2011 Sep 14;12(9):228. doi: 10.1186/gb-2011-12-9-228. Review. PMID: 21920049.

[2] Diagnostic exome sequencing in persons with severe intellectual disability. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Vissers LE. N Engl J Med. 2012 Nov 15;367(20):1921-9. PMID: 23033978.

[3] A post-hoc comparison of the utility of Sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. Neveling K, Feenstra I, Gilissen C, Hoefsloot LH, Kamsteeg EJ, Mensenkamp AR, Rodenburg RJ, Yntema HG, Spruijt L, Vermeer S, Rinne T, van Gassen KL, Bodmer D, Lugtenberg D, de Reuver R, Buijsman W, Derks RC, Wieskamp N, van den Heuvel B, Ligtenberg MJ, Kremer H, Koolen DA, van de Warrenburg BP, Cremers FP, Marcelis CL, Smeitink JA, Wortmann SB, van Zelst-Stams WA, Veltman JA, Brunner HG, Scheffer H, Nelen MR. Hum Mutat. 2013 Dec;34(12):1721-6. PMID: 24123792.

[4] Disease gene identification strategies for exome sequencing. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Eur J Hum Genet. 2012 May;20(5):490-7. doi: 10.1038/ejhg.2011.258. Epub 2012 Jan 18. Review. PMID: 22258526.

Bioinformatic strategies & ontology's

Peter Robinson
Institut für Medizinische Genetik und Humangenetik Charité - Universitätsmedizin Berlin
Germany

Once we have aligned and Q/C'd whole-exome data, and called the variants, the major challenge remains of decided which of the variants might be responsible for the disease we are investigating. Standard pipelines evaluate all variants as to their rarity and known or predicted pathogenicity, and if possible compare results in multiple affected individuals or intersect candidate variants with linkage intervals. Nonetheless, often hundreds of candidate variants remain, and it is often simply not possible to make the diagnosis in a clinical setting or to identify a novel disease gene in research environments. A very active area in bioinformatics research is gene prioritization, that is, ranking candidate genes with respect to their potential relevance to the disease. Compared to previous efforts at prioritization, exome data allows us to intersect an evaluation of the variants with an evaluation of the relevance of candidate genes for the disease.

In this talk, I will present recently published as well as submitted work that employs an analysis of phenotypic similarity or of protein protein interaction networks to prioritize genes. I will begin with a discussion of the Exomiser project (https://www.sanger.ac.uk/resources/databases/exomiser/). The observation that each of our genomes contains about 100 genuine loss-of-function variants makes identification of the causative mutation problematic when using these strategies alone. Here, we propose using the wealth of genotype to phenotype data that already exists from model organism studies to assess the potential impact of these exome variants. We introduce PHenotypic Interpretation of Variants in Exomes (PHIVE), an algorithm that integrates the calculation of phenotype similarity between human diseases and genetically modified mouse models with evaluation of the variants according to allele frequency, pathogenicity, and mode of inheritance approaches in our Exomiser tool. Large-scale validation of PHIVE analysis using 100,000 exomes containing known mutations demonstrated a substantial improvement (up to 54.1-fold) over purely variant-based (frequency and pathogenicity) methods with the correct gene recalled as the top hit in up to 83% of samples, corresponding to an area under the ROC curve of >95%. We conclude that incorporation of phenotype data can play a vital role in translational bioinformatics and propose that exome sequencing projects should systematically capture clinical phenotypes to take advantage of the strategy presented here. More recently we have developed a prioritization framework that takes advantage of a random walk analysis of protein-protein interaction networks. We again showed that this data induces a substantial improvement over purely variant based prioritization. Interestingly, different methods show advantages for different exomes, suggesting that future work show develop a framework for combining different prioritization approaches. In the talk, I will explain the prioritization framework and the algorithms. A workshop will be offered in which we will explore how to use the programs.

As an additional offer to students of this course, it is possible to get a copy of a 150-page handout for a course I am teaching at the Free University of Berlin Bioinformatics Master program on exome sequencing, which includes a lot of practical computer exercises (command line, programming, scripting) for the entire exome pipeline (http://compbio.charite.de/contao/index.php/MedicalGenomics.html). The script is suitable mainly for bioinformaticians, send me a mail if you would like a copy.

Thursday, May 8

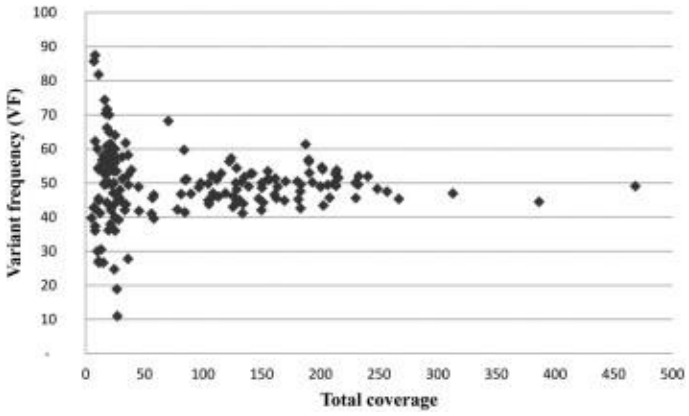
Targeted breast cancer diagnostics

Gert Matthijs
Center For Human Genetics, University of Leuven, Belgium

Hereditary breast cancer can be caused by heterozygous germline mutations in the highly penetrant breast cancer genes, BRCA1 (MIM 113705) and BRCA2 (MIM 600185). Most diagnostic laboratories have moved from direct Sanger sequencing of both genes to mutation screening by NGS. Different NGS platforms are available and have been proposed for BRCA screening. Their properties differ in many ways, e.g. their sequencing capacity, chemistry, and applications.

Multiplexing of amplicons is a way of improving efficiency, especially when there is a large number of amplicons and patients have to be screened. Kits are commercially available, and some of them are CE marked, which facilitates the validation of the method in the diagnostic lab.

When launching an NGS based assay, different parameters have to be tested. For instance, the balance between the coverage and the number of reads with a heterozygous variation has to be examined. They have to be reassessed whenever the pipeline or the method changes. Also, some platforms have difficulties in dealing with the interpretation of homopolymeric stretches. Particular attention has to be paid to platform specific pitfalls.



Distribution of the variant frequency (VF) of heterozygous variations versus the total coverage (from Michils et al. 2012)

Another approach would be to use targeted capture assays. Specifically for breast cancer, Walsh et al (2010) proposed a capture-based method and showed that mutations could be identified in 10 high-risk breast cancer genes. However, when moving from Sanger to amplicon or capture based assays or any other technology, the diagnostic sensitivity for the BRCA1 and BRCA2 genes should not be compromised. Also, it is the responsibility of the community to define the 'core gene' list. This is one of the major issues in choosing the appropriate platform.

As only 10% to 20% of the familial cases are attributable to mutations in the BRCA1 and BRCA2 genes, no mutation is found in at least 80% of the families. Thus, from a diagnostic standpoint, the cost per test remains high. This explains why in most laboratories, a careful selection of patients and families is still made before analysis. This may change as the cost of the analysis will further decrease.

The great majority of the mutations are nonsense mutations and frameshifts. Exonic deletions or duplications have also been reported. There are not easily detected by NGS, and thus, multiplex ligation-dependent probe amplification (MLPA) and other (semi-) quantitative (multiplex) PCR based assays are still being used to complement NGS. The combination of tests that have to be applied, and the order in which they are being offered, has been called the 'diagnostic routing', an important criterion for any type of diagnostic testing.

Finally, for those who are interested in another story: the BRCA1 and BRCA2 genes have been patented, and this led to international actions against gene patenting. The final result is that, in Europe, both genes are still covered by patents, but the original patents have been severely amended. An overview of 10 years of legal actions has been published in Matthijs et al. (2013). In the US, the patenting of genes is now prohibited, as a result of a Supreme Court decision from 2013. It is unclear whether this decision also affects the patenting of diagnostic methods, and whether the link between a disease and a gene defect can still be patented.

References:

Early publications on amplicon-based NGS testing of the BRCA genes

- De Leeneer K, Hellemans J, De Schrijver J, Baetens M, Poppe B, Van Criekinge W, De Paepe A, Coucke P, Claes K. Massive parallel amplicon sequencing of the breast cancer genes BRCA1 and BRCA2: opportunities, challenges, and limitations. Hum Mutat. 2011;32(3):335-44. PMID: 21305653
- Michils G, Hollants S, Dehaspe L, Van Houdt J, Bidet Y, Uhrhammer N, Bignon YJ, Vermeesch JR, Cuppens H, Matthijs G. Molecular analysis of the breast cancer genes BRCA1 and BRCA2 using amplicon-based massive parallel pyrosequencing. J Mol Diagn. 2012;14(6):623-30. PMID: 23034506
- De Leeneer K, De Schrijver J, Clement L, Baetens M, Lefever S, De Keulenaer S, Van Criekinge W, Deforce D, Van Nieuwerburgh F, Bekaert S, Pattyn F, De Wilde B, Coucke P, Vandesompele J, Claes K, Hellemans J. Practical tools to implement massive parallel pyrosequencing of PCR products in next generation molecular diagnostics. PLoS One. 2011;6(9):e25531. PMID: 21980484

BRCA and other panels

- Pritchard CC, Salipante SJ, Koehler K, Smith C, Scroggins S, Wood B, Wu D, Lee MK, Dintzis S, Adey A, Liu Y, Eaton KD, Martins R, Stricker K, Margolin KA, Hoffman N, Churpek JE, Tait JF, King MC, Walsh T. Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. J Mol Diagn. 2014;16(1):56-67. PMID: 24189654
- Walsh T, Lee MK, Casadei S, Thornton AM, Stray SM, Pennil C, Nord AS, Mandell JB, Swisher EM, King MC. Detection of inherited mutations for breast and ovarian cancer using genomic capture

and massively parallel sequencing. Proc Natl Acad Sci U S A. 2010;107(28):12629-33. PMID: 20616022

Standards and guidelines for diagnostic NGS

- Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, Lu F, Lyon E, Voelkerding KV, Zehnbauer BA, Agarwala R, Bennett SF, Chen B, Chin EL, Compton JG, Das S, Farkas DH, Ferber MJ, Funke BH, Furtado MR, Ganova-Raeva LM, Geigenmüller U, Gunselman SJ, Hegde MR, Johnson PL, Kasarskis A, Kulkarni S, Lenk T, Liu CS, Manion M, Manolio TA, Mardis ER, Merker JD, Rajeevan MS, Reese MG, Rehm HL, Simen BB, Yeakley JM, Zook JM, Lubin IM. Assuring the quality of next-generation sequencing in clinical laboratory practice. Nat Biotechnol. 2012;30(11):1033-6. PMID: 23138292
- Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee. ACMG clinical laboratory standards for next-generation sequencing. Genet Med. 2013;15(9):733-47. PMID: 23887774
- Weiss MM1, Van der Zwaag B, Jongbloed JD, Vogel MJ, Brüggenwirth HT, Lekanne Deprez RH, Mook O, Ruivenkamp CA, van Slegtenhorst MA, van den Wijngaard A, Waisfisz Q, Nelen MR, van der Stoep N.Best practice guidelines for the use of next-generation sequencing applications in genome diagnostics: a national collaborative study of Dutch genome diagnostic laboratories. Hum Mutat. 2013;34(10):1313-21. PMID: 23776008

BRCA gene patenting

- Matthijs G, Huys I, Van Overwalle G, Stoppa-Lyonnet D. The European BRCA patent oppositions and appeals: coloring inside the lines. Nat Biotechnol. 2013;31(8):704-10. PMID: 23929344

Exome diagnostics in intellectual disability

Anita Rauch

Anita Rauch, University of Zurich, Institute of Medical Genetics, Schlieren-Zurich, Switzerland

Identification of disease causing mutations in genetically heterogeneous conditions such as intellectual disability by Sanger sequencing is time-consuming, costly and often unsuccessful. The advent of NGS techniques is paving the way for novel large scale approaches with an unforeseen diagnostic power. However, the plethora of variants of unknown significance detected by genome-wide approaches requires distinctive strategies to identify actually disease-related mutations. We recently showed that exome sequencing of patient-parent trios in sporadic cases of unspecific severe intellectual disability may unravel disease causing mutation in more than 50% of previously unsolved cases, with more than 30% of mutations in already known disease genes. Thereby it became also evident, that the current descriptions of phenotypes

associated with mutations in a certain gene, are heavily biased towards certain recognizable patterns. However, while whole exome sequencing may currently provide theoretically the highest cost-efficient diagnostic power, it may miss mutations due to incomplete coverage of certain genes. Therefore in some phenotypes a “clinical exome” limited to a set of genes with currently known monogenic mutations may also be useful.

De novo mutations in human genetic disease

Joris A.Veltman
Department of Human Genetics, Radboud University Nijmegen Medical Centre,
Nijmegen, The Netherlands

Severe disorders with an onset early in life affect fitness, many of them can be considered reproductively lethal. The natural selection against the spread of these disorders in the population should result in a reduction of the occurrence of these disorders, unless these disorders occur de novo by some mechanism. One way in which genetic disorders can arise de novo is by de novo mutations occurring in the germline in each new generation. Unbiased approaches such as exome and genome sequencing, if applied to patients and their unaffected parents, allow us now for the first time to study the presence, frequency and impact of all types of de novo mutations (SNV, indels, CNVs) on genetic disease. In this presentation I will discuss the de novo mutation hypothesis, and recent insight into the role of de novo mutations in rare and common genetic disease. In addition, I will discuss ways to identify the paternal origin of these mutations, and show how these studies have pointed to increased paternal age as a risk factor for disorders caused by de novo mutations.

Recommended reading:

1. Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science 328: 636-9 (2010).
2. Veltman JA and Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet 13: 565-75 (2012).
3. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Jonasdottir A, Wong WS, Sigurdsson G, Walters GB, Steinberg S, Helgason H, Thorleifsson G, Gudbjartsson DF, Helgason A, Magnusson OT, Thorsteinsdottir U, Stefansson K. Rate of de novo mutations and the importance of father's age to disease risk. Nature 488: 471-5 (2012).

NGS and non invasive prenatal diagnosis

Christopher Boustred
North East Thames Regional Genetics Laboratory, Great Ormond Street Hospital NHS Foundation Trust, London

Prenatal testing for fetal aneuploidy and genetic disorders has been available for over four decades (Skirton et al. 2014). In the UK, it is offered to women deemed to have a high-risk pregnancy identified by an abnormal ultrasound scan and / or serum screening result, and those at high prior risk because of a known family history. The purpose of testing is to enable early targeted treatments or surgical intervention and to offer parents the chance to make informed choices about how they wish to proceed with their pregnancy. Furthermore, having the information early on in the pregnancy can help medical staff and parents prepare for the arrival of an affected child.

Traditionally, prenatal testing of these high risk pregnancies is performed on a sample obtained after an invasive test such as chronic villus sampling (CVS) or amniocentesis. Invasive tests have up to a 0.5% risk of miscarriage and can only be performed by qualified clinician at >11 weeks gestation. After the surgical procedure the CVS or amniotic fluid is processed and analysed which can take up to 14 days. Invasive procedures are therefore time consuming and can often cause increased anxiety for parents.

The discovery of cell-free fetal DNA (cffDNA) in maternal plasma (Lo et al. 1997), coupled with the development of low cost next generation DNA sequencing (NGS) technologies (Shendure et al. 2004) has revolutionised prenatal testing (Chitty & Bianchi 2013). CffDNA is DNA derived from the placenta and represents the fetal genome. Approximately 10-20% of the total circulating cell free DNA in the maternal plasma is derived from the placenta (the 'fetal fraction'), although this proportion varies between pregnancies and with gestational age.

In 2008 two proof of principle experiments showed that NGS of cffDNA could accurately detect fetal aneuploidies (Chiu et al. 2008; Fan et al. 2008). The method involved performing NGS on cfDNA and then mapping the millions of short sequence reads to a reference human genome to identifying their chromosome of origin. The number of reads mapped to each chromosome is then counted and compared to a normal 'euploid' set of controls. A significantly higher number of counts mapped to a chromosome is suggestive of a trisomy. Since these early experiments a number of large studies have shown that NGS of maternal plasma can accurately detect the common aneuploidies with high sensitivity and specificity (Palomaki et al. 2011; Bianchi et al. 2012; Norton et al. 2012).

Since 2011 non-invasive prenatal testing (NIPT) for aneuploidy has been available commercially in the USA and there has been a significant reduction in the number of invasive tests being performed. Due to the large potential market for NIPT a number of commercial providers now exist, e.g. Ariosa, Natera, Sequenom, Verinata. Furthermore, large evaluation study projects are under way in many country's public health services.

Regarding prenatal testing for single gene disorders, NGS has significantly reduced the cost of DNA sequencing and the high volume of sequence data produced mean that it is now also possible to expand the breadth of single gene disorders that can be tested for prenatally. However, there has been far less commercial interest in developing these tests because of the smaller market and the requirement to develop bespoke tests that are labour intensive to develop (Lench et al. 2013). Any of these tests in current clinical practice have been developed by academic researchers (Chitty et al. 2011; Chitty et al. 2013).

A number of laboratory and bioinformatic challenges remain before these new approaches can be translated into clinical practice for the more widespread benefit of pregnant women and their families. Challenges include how best to quantify fetal fraction, correct for GC bias, design experiments for single gene disorders and detect of sub chromosomal rearrangements. As well as these challenges, the potential widespread adoption of NIPT poses a number of ethical concerns (Deans et al. 2013; Benn & Chapman 2010)

In this lecture I shall discuss in detail these laboratory, bioinformatic and ethical challenges, and describe how we have begun to implement prenatal testing with NGS in our national health service laboratory.

Benn, P.A. & Chapman, A.R., 2010. Ethical challenges in providing noninvasive prenatal diagnosis. Current opinion in obstetrics & gynecology, 22(2), pp.128–34.

Bianchi, D.W. et al., 2012. Genome-wide fetal aneuploidy detection by maternal plasma DNA sequencing. Obstetrics and gynecology, 119(5), pp.890–901.

Chitty, L.S. et al., 2011. New aids for the non-invasive prenatal diagnosis of achondroplasia: dysmorphic features, charts of fetal size and molecular confirmation using cell-free fetal DNA in maternal plasma. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology, 37, pp.283–289.

Chitty, L.S. et al., 2013. Safe, accurate, prenatal diagnosis of thanatophoric dysplasia using ultrasound and free fetal DNA. Prenatal diagnosis, 33, pp.416–23.

Chitty, L.S. & Bianchi, D.W., 2013. Noninvasive prenatal testing: the paradigm is shifting rapidly. Prenatal diagnosis, 33(6), pp.511–3.

Chiu, R.W.K. et al., 2008. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proceedings of the National Academy of Sciences.

Deans, Z. et al., 2013. Non-invasive prenatal testing for single gene disorders: exploring the ethics. European journal of human genetics : EJHG, 21(7), pp.713–8.

Fan, H.C. et al., 2008. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proceedings of the National Academy of Sciences.

Lench, N. et al., 2013. The clinical implementation of non-invasive prenatal diagnosis for single-gene disorders: challenges and progress made. Prenatal diagnosis, 33, pp.555–62.

Lo, Y.M. et al., 1997. Presence of fetal DNA in maternal plasma and serum. Lancet, 350(9076), pp.485–7.

Norton, M.E. et al., 2012. Non-Invasive Chromosomal Evaluation (NICE) Study: results of a multicenter prospective cohort study for detection of fetal trisomy 21 and trisomy 18. American journal of obstetrics and gynecology, 207(2), pp.137.e1–8.

Palomaki, G.E. et al., 2011. DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. Genetics in medicine: official journal of the American College of Medical Genetics, 13(11), pp.913–920.

Shendure, J. et al., 2004. Advanced sequencing technologies: methods and goals. Nature Reviews. Genetics, 5(5), pp.335–344.

Skirton, H. et al., 2014. Offering prenatal diagnostic tests: European guidelines for clinical practice. European journal of human genetics : EJHG, 22(5), pp.580–6.

Friday, May 9

NGS in population genetics and complex diseases

Paul de Bakker
Departments of Medical Genetics and of Epidemiology
University Medical Center Utrecht
The Netherlands

I will cover a brief history of human disease genetics, starting from Human Genome Project and the HapMap Project through recent genome-wide association studies to interrogate the role of common inherited DNA sequence variation in complex traits and common diseases. I will cover statistical genetic concepts like power, significance (p-value), and illustrate how these differ between common variants and rare variants. With a dedicated focus on the application of next-generation sequencing technology I will also highlight important issues like variant calling, population stratification, and gene-based tests drawing from recent examples in the literature and our own work. A particularly important aspect of my lecture will be the criteria we use in the community for claiming significance. We will discuss how these are fundamentally different between common variant approaches (like GWAS) and rare variant approaches (that employ NGS).

Whole genome sequencing of 3,500 individuals in an isolated population: the Sardinia Sequencing Project

Carlo Sidore
CNR, Institute of Genetics Research, Sassari, Italy

Genome-wide association studies have increasingly furthered our understanding of the molecular basis of many complex traits by finding, through genotyping and imputation, loci associated with many different traits. However, studies based on variants present in common genotyping arrays and imputation panels may not capture the fraction of human genome variation that is rare or geographically restricted and unique to specific populations. To advance our understanding of the genetics of a variety of traits in the Sardinian population, we are studying a sample of 6,602 individuals recruited from the population of a cluster of 4 small towns in Sardinia. Using whole genome sequencing, we sequenced DNAs from 3,514 Sardinian individuals enrolled either in this project or in a parallel project on autoimmune diseases, at an average depth of coverage of ~4X. We successfully identified and genotyped >23M single nucleotide polymorphisms (30.6% of them novel, not in dbSNP v135) with an error rate of 0.15%. To increase the power to detect

association, we are using the haplotypes generated by sequencing of these individuals to impute missing genotypes in the remaining 6,602 already genotyped with an integrated map of 4 GWAS Illumina arrays (ImmunoChip, CardioMetaboChip, ExomeChip and OmniExpress). Strikingly, imputation using our Sardinian reference panel shows greatly increased accuracy when compared to an equal size reference panel of European haplotypes generated by the 1000 Genomes Project (average imputation accuracy, rsqr=0.9 compared to 0.75 for alleles with frequency 1-3%). With a larger reference panel, imputation accuracy of variants with frequency 1-3% reaches 0.94 %, giving us the possibility of analyzing the rare frequency domain in the Sardinian population.

As an example of the advantages of analyzing population specific rare variation, we will discuss the Q39X mutation in the *HBB* gene, which is common in Sardinia (MAF ~5%) but very rare elsewhere. The variant is associated with a variety of blood phenotypes. For LDL cholesterol, this variant has the second largest effect among the variants previously found with standard GWAS analysis. Our approach thus increases the power of detecting population specific association.

Cancer genome sequencing

Ian Tomlinson
Wellcome Trust Centre for Human Genetics,
University of Oxford, UK

The ability to sequence cancer genomes naturally provides the opportunity to achieve a much greater understanding of how cancers develop in terms of their dynamics, clonal structure and mutational processes. It also allows the discovery of new driver genes and provides the ability to test a number of specific hypotheses. In this presentation, I use our own whole-genome sequence data from a modest number of bladder cancers to illustrate the opportunities and challenges associated with cancer genome sequencing. I compare the virtues of whole genome and exome sequencing and the utility of ancillary data such as those derived from SNP arrays or methylation profiling. Finally, I examine the utility of sequencing focussed panels of cancer genes, such as those that are beginning to be used in molecular diagnostics. A great many methodological challenges remain, and this is especially true of the great majority of the genome for which specific functions remain unidentified. I conclude with a brief tour of some of the changes in our us understanding that cancer genome or exome sequencing has already brought, including clonal structures, mutation signatures, new mutational processes, new driver genes such as those involved in controlling chromatin structure, and the apparently paradoxical nature of some changes in which the same pathway is activated in one cancer type and inactivated in another.

Selected references:

1: Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MD, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ, Ding L. Mutational landscape and significance across 12 major cancer types. Nature. 2013 Oct 17;502(7471):333-9.

2: Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013 Jul 4;499(7456):43-9.

3: Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012 Oct 4;490(7418):61-70.

4: Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012 Jul 18;487(7407):330-7.

5: Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014 Jan 23;505(7484):495-501.

6: Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandoth C, Reimand J, Lawrence MS, Getz G, Bader GD, Ding L, Lopez-Bigas N. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Sci Rep. 2013 Oct 2;3:2650.

7: Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jäger N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, López-Otín C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdés-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature. 2013 Aug 22;500(7463):415-21.

8: Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, Yates LR, Papaemmanuil E, Beare D, Butler A, Cheverton A, Gamble J, Hinton J, Jia M, Jayakumar A, Jones D, Latimer C, Lau KW, McLaren S, McBride DJ, Menzies A, Mudie L, Raine K, Rad R, Chapman MS, Teague J, Easton D, Langerød A; Oslo Breast Cancer Consortium (OSBREAC), Lee MT, Shen CY, Tee BT, Huimin BW, Broeks A, Vargas AC, Turashvili G, Martens J, Fatima A, Miron P, Chin SF, Thomas G, Boyault S, Mariani O, Lakhani SR, van de Vijver M, van 't Veer L, Foekens J, Desmedt C, Sotiriou C, Tutt A, Caldas C, Reis-Filho JS, Aparicio SA, Salomon AV, Børresen-Dale AL, Richardson AL, Campbell PJ, Futreal PA, Stratton MR. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012 May 16;486(7403):400-4.

9: Cazier J-B, Rao SR, Mclean CM, Walker AL, Wright BJ, Jaeger EEM, Kartsonaki C, Marsden L, Yau C, Camps C, Kaizaki P, WGS500 Consortium, Taylor J, Catto JW, Tomlinson IPM, Kiltie AE, Hamdy FC. Nat Comms. 2014; in press.

Exome sequencing to study rare and common variation in diabetes

Amelie Bonnefond
CNRS UMR8199, Lille Pasteur Institute, Lille II University, Lille, France

Type 2 diabetes (T2D) is a complex disorder that has developed into a major health problem responsible for early morbidities and mortality with the worldwide prevalence doubling every 15 years. Although T2D is considered preventable, it is usually not reversible and current medications are generally unable to stop disease progression, resulting in poor glucose control and degenerative complications. T2D is characterized by insulin secretory dysfunction of pancreatic β-cells combined with insulin resistance, but the primary mechanisms are still largely debated. A genetic contribution is well recognized in the diverse forms of both early-onset (mostly known as monogenic) and adult-onset diabetes. Through several subsequent genetic approaches, we know that these diverse forms of diabetes are characterized by vast genetic heterogeneity, highlighting the complexity of the signalling pathways that are pivotal for the biology of the pancreatic β-cell. Although recent genetic discoveries have provided new insights into β-cell physiopathology (via genome-wide association studies for polygenic forms of T2D and candidate gene studies for monogenic forms of diabetes), major gaps in our understanding of the genetic basis of diabetes remain; a proportion of monogenetic diabetes cases remain to be explained and the heritability of common T2D is much higher than the overall impact of the known genetic variants. Thus, a substantial genetic 'dark matter' remains to be elucidated.

Next-generation sequencing, including whole-exome sequencing (WES), has brought new fantastic expectation to elucidate most of this dark matter. WES has been quite successful in elucidating novel monogenic forms of diabetes (including neonatal diabetes associated with pancreatic agenesis or maturity-onset diabetes of the young [MODY]), and new susceptibility genes for T2D, even if the community seems to be quite disappointed. I will discuss about the strength and the limit of WES in both monogenic and polygenic forms of diabetes.

Furthermore I will discuss about the breakthrough of next-generation sequencing into highly sensitive molecular diagnosis of known monogenic forms of diabetes.

Saturday, May 10

Analysis of mitochondrial genomes

Adam Ameur
Science for Life Laboratory, National Genomics Infrastructure,
Uppsala University, Sweden

Mitochondria are organelles that are crucial for the energy production in eukaryotic cells, and mitochondrial defects have been implicated in a wide range of degenerative diseases, aging and cancer[1]. The number of mitochondria in a human cell varies between one and several thousand, with large differences between tissues. Each mitochondrion carries its own DNA sequence. The human mitochondrial DNA (mtDNA) is a small circular molecule consisting of approximately 16,600 bases. Because of the small size of the mtDNA sequence, and other unique properties such as maternal inheritance and elevated mutation rate, many studies of human evolution are based on mtDNA analysis.

With the emergence of next-generation sequencing (NGS) technologies it is now possible to study the mtDNA sequence at a new level of resolution. Previously, using Sanger's sequencing, studies were practically limited to generating a consensus mtDNA sequence for a given sample. With NGS technologies it is now possible to see also low frequency mutations present only in a small fraction of the mitochondria in a cell[2]. It has been shown that NGS can detect mutations that are present in as few as one out of 10,000 mitochondrial genomes[3]. This opens up new possibilities for studying accumulation of mitochondrial DNA mutations during disease progression and aging[4].

Analysis of mtDNA sequencing data requires specific considerations that are not always handled by tools used for 'normal' NGS data. Since the mitochondrial genome is very small, the sequencing instruments will typically generate an extremely high coverage of the mtDNA sequence, in some cases up to 100,000X coverage or more. Many NGS tools cannot handle that type of extreme coverage, something that can make the analysis challenging.

The future for mtDNA sequencing looks very interesting. With the development of so-called third generation sequencing instruments capable of generating very long reads, it might soon be possible to obtain a full-length mitochondrial genome in a single high quality read. This would enable us to study complete mitochondrial genomes on a single molecule scale.

1 Wallace, D. C. Mitochondrial diseases in man and mouse. *Science* **283**, 1482-1488 (1999).

2 He, Y. *et al.* Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* **464**, 610-614, doi:10.1038/nature08802 (2010).

3 Ameur, A. *et al.* Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS genetics* **7**, e1002028, doi:10.1371/journal.pgen.1002028 (2011).

4 Park, C. B. & Larsson, N. G. Mitochondrial DNA mutations in disease and aging. *The Journal of cell biology* **193**, 809-818, doi:10.1083/jcb.201010024 (2011).

Retrieval and annotation of human mitochondrial mutations from NGS experiments

Francesco Calabrese

CEINGE (Centro Ingegneria Genetica) Advanced Biotechnologies, Naples, Italy and Dip. Bioscienze, Biotecnologie e Biofarmaceutica, Univ. Of Bari

With the advent of NGS techniques a huge quantity of genomic data can benow used for genetic and pathological studies.New specific pipelines [1,2] allow the extraction of mitochondrial reads from experiments not exactly designed for this target, where library preparation platforms that enrich for DNA coding sequence show different capture efficiencies. The yielded variable numbers of off-target data from human NGS can be used to assemble mitochondrial genomes, taking into account the cross-hybridization ofmtDNA with baits that overlap Nuclear mitochondrial Sequences (NumtS).

The expected high coverage is one of the useful properties which helps clinicians in solving one of the most important mitochondrial problematic features: the heteroplasmy.

Further improvements of the above mentioned methods and the development of new mitochondrial databasesallow the estimate of pathologic variants together with other mitochondrial data such as variability, haplogroup assignment, pathogenicity scoresand the somatic status.

A practical example of the application of highly automated bioinformatics pipelineonNeuroblastomas shows the biological considerations underlying thisargument.

[1] Mitochondrial genomes gleaned from human whole-exome sequencing.
Picardi E, Pesole G. Nat Methods. 2012 May 30.

[2] MToolBox (http://sourceforge.net/projects/mtoolbox/)

Posters of Students

Two Genes Associated with Hearing Loss in a Consanguineous Family Are Identified by Next-Generation Sequencing

Nada D. Farran1,2, Zippora Brownstein2, Ofer Yizhar-Barnea2, Karen B. Avraham2, Stavit Shalev1

1Genetics Institute, Ha'Emek Medical Center, Afula, Israel and Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; **2**Department of Human Molecular Genetics and Biochemistry, Sackler Faclty of Medicine, Tel Aviv University, Tel Aviv, Israel

Background: The first gene for autosomal recessive hearing loss, GJB2, was identified in 1997. Today hundreds of mutations in more than 80 genes are known to be associated with hearing loss. A significant acceleration in deafness-gene discovery has been made with Next-Generation Sequencing (NGS). Nevertheless, a significant portion of hereditary hearing loss remains unsolved. This is particularly true of the Middle Eastern population, with many different ethnic groups and high rates of consanguinity. We performed a search for the genetic basis for hearing loss in a large family with several consanguineous marriages, Family H.

Methods: First, we performed NGS using a targeted genomic capture approach with 284 genes, including 121 human genes and 163 human orthologous of mouse deafness genes, in 150 deaf probands, including Family H. Next, three members of Family H underwent whole exome sequencing (WES). Segregating variants were validated by Sanger sequencing.

Results: Targeted genomic capture and NGS resulted in the doubling of the number of deafness genes in our Middle Eastern cohort. However, this approach did not uncover the genetic basis for deafness in Family H. WES led to the discovery of two causative mutations in two different genes, OTOF and SLC25A21, in different branches of the family. The OTOF mutation was previously found to lead to auditory neuropathy. Subsequent clinical analysis of this branch of Family H determined that they too have auditory neuropathy. The second branch harbored a mutation in SLC25A21, a mitochondrial transporter. Functional assays are ongoing to determine the pathogenicity of the SLC25A21 mutation and its contribution to hearing impairment.

Conclusions: The results demonstrating that two different genes might be responsible for hearing loss in the same family with high rates of consanguinity emphasizes the utility of WES for resolving the genetic basis of hearing loss. However, this search also highlights the complexity of determining the causative mutation. Functional assays to evaluate the defective proteins will provide precise conclusions.

Gene set enrichment analysis highlights the relationships of the NF-κB/Snail/YY1/RKIP circuitry genes in multiple myeloma.

Apostolos Zaravinos and Qiang Pan-Hammarström
Department of Laboratory Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

Presenting author's e-mail contact: apostolos.zaravinos@ki.se

Background: Multiple myeloma (MM) is characterized by the clonal proliferation of plasma cells in the bone marrow, usually with elevated serum and urine monoclonal paraproteins and associated end organ damage. The presence of a dysregulated NF-κB/Snail/YY1/RKIP loop was recently established in metastatic prostate cancer cells and non-Hodgkin's lymphoma; however, its involvement in MM has yet to be investigated. Furthermore, not much is known regarding the genes that are correlated with this circuitry.

Objectives: Aim of the study was to investigate the role of the NF-κB/Snail/YY1/RKIP circuitry in MM and how each gene is correlated with the remaining genes of the loop.

Methods: Using GSEA and Gene Neighbors Analysis in data received from four datasets included in the Multiple Myeloma Genomics Portal (MMGP) of the Multiple Myeloma Research Consortium (MMRC), we identified various enriched gene sets associated with each member of the NF-κB/Snail/YY1/RKIP circuitry. The publicly available Gene Expression Omnibus (GEO) datasets GDS2643, GSE24870, GSE27838, GSE2113 and GSE24990 were also analyzed for the expression levels of the circuitry genes.

Results: In each dataset, the 20 most co-expressed genes with the circuitry genes were isolated subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment. Among them, we highlighted the genes FNDC3B, TPD52, MBNL1, MFAP2 and BRD4. The FOXO4, GATA binding factor, Sp1 and AP4 were the transcription factors that most likely affect the expression of the NF-κB/Snail/YY1/RKIP circuitry genes. GEO datasets computational analysis revealed elevated YY1 and RKIP levels in MM vs. the normal plasma cells, as well as elevated RKIP levels in MM vs. normal B lymphocytes.

Conclusion: The present study highlights the relationships of the NF-κB/Snail/YY1/RKIP circuitry genes with specific cancer-related gene sets among four MM datasets.

EHLERS DANLOS SYNDROME: A NEW CUSTOM PANEL BY TRUE SEQ CUSTOM AMPLICON APPROACH

Francesca Cortini

Medical Genetics Laboratory, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Italy and Center of Molecular and genetic Epidemiology , Dept. Of Environmental and Occupational Health, Università di Milano, Fondazione Ca Granda, IRCCS Ospedale Maggiore Policlinico, Milan, Italy

INTRODUCTION

Over the last few years, Next Generation Sequencing (NGS), has dramatically changed approaches to basic, applied and clinical research. NGS combines different strategies that are dependent on a combination of template preparation, sequencing and genomic allignment and allows the simultaneously analysis of hundred thousand DNA fragments.

The vastest application of NGS may be resequencing of human genetics to enhance our understanding of how genetic differences affect health and disease.

AIM OF THE PROJECT

For our porject, by Illumina Platform/MiSeq System, we are going to perform target resequencing of the coding region and 5' and 3' UTR of specific genes associated to Ehlers Danlos syndrome (EDS) with a specific kit from Illumina Technologies to identify the presence of the single nucleotide variations.

PLAN OF EXPERIMENT

By True Seq Custom Amplicon we plan a custom panel composed of 23 genes. In base of the most rappresentative phenotypes on our EDS database, we choose characteristic collagen and extra cellular matrice genes. Custom panel would rapresent a good and fast method to identify new mutations and for genetic counselling.

Identification of a Novel Locus for Arrhythmogenic Right Ventricular Cardiomyopathy on chromosome 19p13.3

Li Mura I.1, Poloni G.1*, Bauce B.2, Vazza G.1, Rigato I.2, Mazzotti E.2, Calore M.1, Lorenzon A.1, De Bortoli M.1, Daliento L.2, Basso C.2, Corrado D.2, Thiene G.2, Rampazzo A.1*

[1]Department of Biology, University of Padua, Padua, Italy
[2]Department of Cardiac, Thoracic and Vascular Sciences, University of Padua, Padua, Italy
*equally contributed

Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC) is an autosomal dominant cardiomyopathy characterized by myocardial atrophy, fibrofatty replacement and fibrosis, which mainly involves the right ventricle. It is recognized worldwide as the second most common cause of unexpected sudden death among young people and athletes. Thirteen causative genes have been identified, with a central role of the desmosomal genes. Only 30% to 50% of ARVC patients resulted to carry desmosomal gene mutations, suggesting the involvement of other disease genes.

Here we report the identification of a novel ARVC locus in a large family without any mutations in desmosomal genes. A number of 45 family members were genotyped using the High−density SNP arrays, which includes 370,404 SNP markers with average spacing of 7.9 Kb. Linkage analysis, using an affected-only approach, was performed and allowed us to exclude linkage of ARVC to markers in 86% of the genome and to identify a unique, significant linkage peak on chromosome 19p13.3 (pLOD=3.85). CNVs analysis excluded the presence of rare genomic rearrangements. Among the 44 genes mapped on the predicted locus, *PTPRS* gene was selected and directly sequenced as a candidate gene, but only common SNPs were identified. Therefore, exome sequencing on HiSeq2000 platform was performed for two affected family members, after exome capture with SureSelect 50 Mb kit. Within the critical region (2Mb), 32 variants shared by both individuals were identified, but no one likely represents a new mutation variant. Our data suggest the existence of a novel ARVC locus on chr19p13.3 and further investigation might reveal new causative mutations.

Unraveling oligogenic inheritance and mutational load in detrmining Joubert Syndrome phenotype

Giuliano Bolondi

Joubert Syndrome (JS) is an autosomal recessive or X-linked condition presenting in neonatal age with hypotonia, ataxia, abnormal ocular movements, psychomotor delay and cognitive impairment and variable multiorgan involvement.Magnetic resonance imaging shows a peculiar cerebellar and brainstem malformation, the so-called “Molar Tooth Sign” (MTS), which is mandatory for the diagnosis of JS.

So far, 24 causative genes have been identified, that overall account for only ~50% cases. However, their overall mutational load and phenotype correlates remain largely unknown. All JS causative genes encode proteins of the primary cilium, a nearly ubiquitous organelle playing key functions in embryonic and adult tissues,making JSpart of the expanding group of “ciliopathies”. These are characterized by great clinical and genetic heterogeneity: indeed, the same phenotype can be caused by mutations in distinct genes, and conversely the same gene can be responsible for different ciliopathies.1 Such variability support an oligogenic model of inheritance and suggests the existence of genetic modifiers.

In our center, we have recruited over 400 JS families, of whom detailed clinical features and brain imaging were systematically collected. DNA samples from 300probands were adequate for next generation sequencing (NGS) studies, including 68 patients carrying mutations in known genes. This cohort will undergo NGS-based target resequencing(Solid 5500xL platform) of 120 genes known to be causative for

ciliopathies or implicated in the ciliary formation or function. The genes have beendistributed ontotwodistinct panels, designed to cover all coding exons and the flanking 50bp, as well as untranslated regions of each gene, for a total of about 400 kb genomic sequence.

The rationale to adopt target sequencing instead of whole exome sequencing as a first pass screening of our whole cohort is based on the cost-effectiveness of the technique; most importantly, the much lower number of tested genes is expected to lead to a significant increase of coverage on target sequences (we aim at a minimal 20x coverage per single base). This will minimize false negative results and will also prevent detection, validation, and interpretation of countless variants in genes possibly unrelated to the patient's phenotype, as expected using a exome approach.

The BioinformaticUnit at our Institution will perform data analysis, based on a quality consensus-based workflow, which will start mapping short-reads against a known reference genome using the Lifescope software. Then, it will select the best alignment according to mapping and site-specific quality scores. To confer higher confidence, both short-reads and alignments will undergo an extensive evaluation of the base-per-position and mapping-per-read quality scores. This will drive to define an optimal filtering procedure. Additionally, an exhaustive screening of the perbase/per-exone coverage will be given, so as to provide with a preliminary insight of the potential weaker genomics regions. Alignments will be scanned by three SNPs and short-Indels callers (DiBayes, the GATK and SAMtools).

Genotypes will be matched and scored according to the degree of consensus among the three tools, and polymorphisms will be annotated using distinct tools, as appropriate (ANNOVAR, snpEff, dbSNP, 1000 Genomes, EVS).

Finally, the dangerousness of each variant will be checked by a conservational and structurally conformational inspection by PolyPhen-2, Sift and MutPred software.2

Statistical analysis will be performed in collaboration with our Unit of Biostatistics. The different clinical phenotypes will be considered as outcomes and analyzed using a Random Forest (RF) approach, a powerful machine learning statistical algorithm, to obtain reliable results using a relative small sized sample. RF is an ensemble of classification trees.It will be applied to classify subjects according to their clinical phenotype using the whole set of covariates (i.e. genes and clinical-pathological characteristics). The utility of RFs stems not only from their classification ability, but also from the information they yield about the structure of the model underlying the data. The latter attribute can be particularly appealing in problems in which the number of predictors is large and their interactions may occur, as it is the case of this project. Analyses will be performed using Random Forest package implemented in R.3

We aim to define the mutational load and phenotypic spectrum of known JSRD genes, in order to determine genephenotype (and, in selected cases, genotype-phenotype) correlates, to search for oligogenic inheritance, and to identify possible genetic modifiers of the phenotype. This study will determine the prevalence and phenotypic spectrum of JS-related mutations and the role of modifier genetic variants in a large cohort of patients, and help understand the pathogenetic mechanisms leading to cerebellar derangement during embryonic development. In turn, this will help develop efficient algorithms for pre- and post-natal diagnosis, improve genetic counseling and help treat or delay the progression of associated conditions such as renal failure.

1. Romani, M., Micalizzi, A. & Valente, E. M. Joubert syndrome: congenital cerebellar ataxia with the molar tooth. *Lancet Neurol.*12, 894–905 (2013).
2. Castellana, S. & Mazza, T. Congruency in the prediction of pathogenic missense mutations: state-of-the-art
web-based tools. *Brief. Bioinform.*14, 448–59 (2013).
3. Chen, X. & Ishwaran, H. Random forests for genomic data analysis. *Genomics*99, 323–9 (2012).

The identification of possible causative genetic variants in two siblings with idiopathic hearing loss using exome sequencing

Jernej Kovač1, Gašper Klančar1, Saba Battelino2, Katarina Trebušak Podkrajšek1

[1] Center for Medical Genetics, University Children's Hospital, UMC Ljubljana, Slovenia
[2] Clinic for Otorhinolaryngology and Cervicofacial Surgery, UMC Ljubljana, Slovenia

Sensorineural hearing loss (SNHL) presents one of the most common disabilities, with severe impact on human social and economic status. More than 100 genome loci and around 40 genes are known to be involved in the aetiology of SNLH. The next generation high-throughput sequencing (NGS) facilitate the identification of potential genetic variants and their role in the development of SNLH.

We are presenting two siblings with severe hearing loss and preserved ocular function. Both patients, male and female, had almost identical clinical picture with hearing loss in the early childhood. Their hearing was partially restored with cochlear implant surgical procedure. On the other hand, they had their vision perfectly preserved excluding the possibility of having Usher syndrome.

The exome sequencing (TruSight One, Illumina) of both siblings was performed to identify possible underlying genetic causes of the SNHL. Due to the almost identical development of SNHL we assumed that shared genetic variants are the cause of the SNHL. The family segregation of identified variants was evaluated with Sanger sequencing.

Approximately 1260 genetic variants for each participant were identified with sufficient sequence quality to include them into the further analysis. Reducing the number of possible causative variants trough implementation of different queries (sequencing depth, MAF, consequence, etc.) four variants were identified as possible cause for SNHL. Both siblings are compound heterozygotes for c.6847G>A (p.Val2283Ile) and c.7762G>C (p.Glu2588Gln) variant in the gene for cadherin 23 (*CDH23*). The carrier of c.6847G>A was a father and the carrier of c.7762G>C was a mother. Both siblings had additional but different mutation present in the gene for fascin 2 (*FSCN2*). The male sibling had the heterozygous stop gained mutation c.130C>T (p.Gln44Ter) while the female sibling is the carrier of the heterozygous deleterious missense mutation c.412C>T (p.His138Tyr). The origin of the *FSCN2* mutations is still under investigation.

Cadherin 23 as well as fascin 2 are involved in the function of stereocilia filaments which enable the transformation of the sound into neural signals. The combination of *CDH23* and *FSCN2* genetic variants is reported to be favourable for increased rate of hearing loss. While the presence of genetic variants in the *CDH23* resulting in complete protein function loss may cause the development of Usher syndrome (loss of

hearing and vision), the variants that at least partially preserve protein function result in non-syndromatic hearing loss with preserved ocular function. Additionally, the simultaneous presence of mutations in *FSCN2* is associated with accelerated deterioration of hearing in the carriers of *CDH23* mutations.

The NGS technology is rapidly transforming the field of genetic medical testing and diagnostics, enabling fast and reliable acquisition of genetic data, reducing the time to identify the underlying genetic causes of specific ailments.

Targeted resequencing approach to investigate the mutational landscape associated to platinum resistance in EOC

Mariacristina Di Marino[1], Luca Beltrame[1], Lara Paracchini[1], Robert Fruscio[2], Vittoria Fotia[3], Tommaso Grassi[2], Luca Clivio[1], Maurizio D'Incalci[1] and Sergio Marchini[1]

Introduction Despite initial response to first line platinum-based chemotherapy, more than 80% of high grade serous ovarian cancer patients relapse and develop resistance. The molecular and genetic features involved in drug resistance are still unknown. By gene expression profile in a cohort of patients from which matched biopsies were taken at primary surgery (PS-O) when tumor was sensitive to chemotherapy and at time of relapse (SCR) when the tumor was resistant, we identified EMT pathway as a key player in tumor relapse (Marchini et al., 2013). Here we investigate the genomic alterations driving drug resistance by performing targeted DNA resequencing on our cohort of SCR and PS-O samples.

Methods DNA libraries enriched in a selected panel of 30 genes encompassing key players of signal transduction, cell cycle and DNA repair were generated using TruSeq Custom Amplicon kit and sequenced on MiSeq (Illumina). Data were analyzed using a high performance cluster computing platform (Cloud4CARE project).

Results Analysis identified a total of 166 mutations (152 SNPs and 14 InDels), of which 51 affecting PS-O and 115 SCR. With a 500X coverage, we observed BRCA1, BRCA2 and TP53 mutated in the majority of cases. In addition, the PI3K pathway was found mutated in SCR samples only.

Conclusions Our preliminary results suggest two main conclusions:

1- genomic alterations (SNPs or InDels) were more frequent in SCR compared to PS-O;

2- most of the mutations affected genes belonging to DNA and PI3K pathway.

1. Department of Oncology, IRCCS-Istituto di Ricerche Farmacologiche 'Mario Negri', Milano, Italy.
2. Clinic of Obstetrics and Gynecology, University of Milano-Bicocca, San Gerardo Hospital, Monza, Italy.
3. Medical Oncology Unit, IRCSS- Fondazione Salvatore Maugeri, Pavia, Italy.

Targeted sequencing of mitochondrial exome in pediatric patients with mitochondrial diseases.

Markéta Tesařová[1], Daniela Záhoráková[1], Tomáš Honzík[1], Hana Kratochvílová[1], Alžběta Vondráčková[1], Viktor Stránecký[2], Marie Rodinová[1], Hana Hansíková[1], Jiří Zeman[1],

[1]Department of Pediatrics and Adolescent Medicine and [2]Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Czech Republic

Mitochondrial diseases (MD) represent both clinically and genetically heterogeneous group of disorders. More than 130 nuclear genes have been described so far whose mutations lead to MD. Accurate targeting of the genetic analysis based on clinical symptoms or laboratory tests is possible for only a few types of MD (MNGIE syndrome, mitochondrial encephalo-cardiomyopathy - *TMEM70*, Leigh syndrome due to *SURF1* mutations). For most of MD, even specialized enzymatic and protein analyses do not allow unambiguously to narrow a group of candidate genes. Next generation sequencing is a significant milestone in discovery of genetic bases of inherited diseases and it already transforms genetic diagnostics. In last two years, more than 30 MD-genes were identified by exome sequencing. In a group of 48 patients with mitochondrial disease we performed targeted sequencing of mitochondrial exome (1233 nuclear encoded genes mostly based on MitoCarta Inventory; http://www.broadinstitute.org/pubs/MitoCarta/index.html). In 10 patients, mutations in known MD-genes *TSFM, COX10, AIFM1, TK2, MGME1, NDUFA1, NDUFA10, TRMU* were found. In 6 patients candidate variants in genes with previously unknown association with MD were selected. All prioritized variants are further evaluated to support their pathogenicity. *Supported by research projects GAČR 14-36804G, RVO-VFN64165/2012 and grants IGA NT13114/4, IGA NT13120/4.*

The use of Whole Exome Sequencing to identify new phenotypes of known genetic syndromes

Weisz Hubshman M[1,2] ,Tzur S[3] ,Behar DM[4] ,Smirin Yosef P[5], Basel-Vanagaite L[1,2,5,6]

[1]Pediatric Genetics, Schneider Children's Medical Center of Israel, Petach Tikva. [2]Raphael Recanati Genetic Institute; [3]Molecular Medicine Laboratory, Rambam Health Care Campus, Haifa, Israel. [4]Genetic Institute, Rambam Health Care Campus, Haifa, Israel. [5]Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel. [6]Felsenstein Medical Research Center, Rabin Medical Center, Beilinson Campus, Petach Tikva

The clinical diagnosis of genetic syndromes is getting complicated with the recognition of novel genetic syndromes with variable expressivity. In this study we demonstrate the use of Whole Exome Sequencing (WES) in identification of atypical clinical presentation of a known genetic syndrome.

We investigated a six years old girl with severe developmental delay, failure to thrive, microcephaly and dysmorphic features. Using WES a novel missense mutation in the MLL gene was identified. This gene was previously shown to be involved in Wiedemann-Steiner syndrome. The characteristics of this syndrome

include developmental delay, mild to moderate growth problems, hirsutism particularly at the elbows, dysmorphic features and mild limb abnormalities.

This study shows that WES is useful in the recognition of rare or atypical phenotypes of known syndromes which cannot be identified clinically. Our case shows a severe phenotype of Wiedemann-Steiner syndrome and our hypothesis is that in contrast to the frameshift mutations causing premature termination of protein identified as causative of this syndrome, missense mutation can cause dominant negative effect or gain of function and as a result of this a much more severe phenotype.

Identification of putative precursor lesion of papillary thyroid carcinoma by cyclin D1 overexpression and p38 MAPK phosphorylation

Monika Lamba Saini, Birgit Weynand #, Jacques Rahier*, Michel Mourad†, Marc Hamoir††,*

*Etienne Marbaix**

*Service d'anatomie pathologique, Cliniques universitaires Saint-Luc, Université catholique de Louvain, B-1200 Bruxelles, Belgium

Service d'anatomie pathologique, CHU UCL Mont Godinne, Yvoir 5530, Belgium

† Service de transplantation rénale et chirurgie endocrinologique, Cliniques universitaires Saint-Luc, Université catholique de Louvain, B-1200 Bruxelles, Belgium

†† Service d'oto-rhino-laryngologie et chirurgie cervico-faciale, Cliniques universitaires Saint-Luc, Université catholique de Louvain, B-1200 Bruxelles, Belgium

Corresponding author's address and e-mail:

Dr. Monika Lamba Saini, MD, DNB
Anatomie pathologique
Cliniques universitaires Saint- Luc
Université catholique de Louvain
Avenue Hippocrate, 10/T-1,
Brussels 1200,
Belgium
Phone: +32-02-764-6751
E-mail: monikalamba@gmail.com, monika.lambasaini@uclouvain.be

Introduction: Papillary thyroid cancer (PTC) is the commonest endocrine malignancy (1). Though significant progress has been made to understand the pathways involved in the tumorigenesis of PTC, no precursor lesion has been identified yet. PTC also frequently carries several genetic alterations in genes

coding for proteins that activate the mitogen-activated protein kinases (MAPK) signaling pathway, which plays a key role in the regulation of cell growth and differentiation (2).

Aims:

- To evaluate the genetic abnormalities leading to cyclin D1 overexpression in follicular variant of PTC arising in a follicular adenoma and also to elucidate the premalignant nature, if any, of this lesion.

- To investigate the MAPK signalling pathway [extracellular-regulated kinase (ERK), Jun N-terminal Kinase (JNK) and p38] involved in tumorigenesis of PTC by using IHC and Western blotting.

Materials and Methods:

Thirteen cases of metastatic PTC, papillary microcarcinoma and follicular variant of PTC (FVPTC) were identified from a histological review of 510 cases. In addition, 13 cases of a subset of follicular adenomatoid nodules with focal areas showing nuclear features characteristic of PTC, identified as putative PTC precursor lesion, were also analyzed. Immunohistochemical analysis of galectin-3, HBME-1, CK 19 and the proliferation markers Ki 67 and cyclin D1 was performed. Lesions were analyzed for cyclin D1 gene amplification by fluorescent in-situ hybridization.

Twenty samples of PTC and its follicular variant (11 classic PTC, 9 follicular variant of PTC) were retrieved from the UCL biolibrary. Frozen sections of samples and normal thyroid tissue (control) from the biobank were solubilized to perform Western blots of total and phosphorylated forms of ERK, JNK and p38. Corresponding paraffin blocks were retrieved and IHC analysis of total and phosphorylated forms of IHC for ERK, JNK and p38 was performed.

Results:

All putative precursor lesions showed immunolabelling of cyclin D1, Ki 67; 11/ 13 cases showed immunolabelling of CK 19; 10/13 cases showed immunolabelling of HBME-1 and 4/13 cases showed immunolabelling of galectin-3. Surrounding adenomatoid areas showed no to faint focal staining in all thirteen cases of cyclin D1, HBME-1 and galectin-3. A low rate of cyclin D1 gene amplification was identified in a significant proportion of cells in the putative precursor lesion as compared to surrounding benign adenomatoid areas.

ERK activation was seen as nuclear immunolabelling in less than 10 % of tumour cells in 15/20 PTC cases. JNK activation was seen as nuclear and membranous immunolabelling in less than 10 % of tumour cells in 7/20 cases. However, total ERK and total JNK were seen as nuclear and cytoplasmic immunolabelling in more than 50% of tumour cells in 10/20 and 16/20 cases respectively. In contrast, p38 MAPK phosphorylation was seen as abundant nuclear and cytoplasmic immunolabelling in 11/20 cases (6 FVPTC and 5 classical PTC. Total p38 was immunolabelled in 14/20 cases. A one way ANOVA test showed significant difference between the ERK, JNK and p38 phosphorylation ($p<0.01$). By Western blotting, p38 was detected in all twenty PTC samples and its phosphorylated form was detected in 17/20 samples. The signals were much stronger in the follicular variants of PTC as compared to classical PTC. p38 was also detected in normal thyroid tissue but not its phosphorylated form.

Conclusions:

Increased expression of cyclin D1 and amplification of its gene along with immunolabelling of HBME-1 in areas showing cytological features of PTC within follicular adenomatoid nodules suggest that these areas could correspond to a precursor lesion of follicular variant of PTC.

There is variable expression and activation of the MAPK pathways in PTC but p38 appears to be activated in a larger proportion of PTC than ERK or JNK. The role of p38 in tumorigenesis of PTC and its variable expression in different variants of PTC is being further investigated by the effect of specific inhibitors of the various MAPK pathways on cellular growth assays.

Future perspectives:

We aim to identify and analyze novel molecular species in classical PTC, follicular variant of PTC (FVPTC) and follicular thyroid carcinoma (FTC) by using ultra-high throughput sequencing technology, and to investigate whether they are related to the MAP kinase signal transduction pathway by validation of the identified potential biomarkers by RT-PCR and IHC on tissue microarrays. A particular focus of the study will be to use the molecular species identified in the transcriptome profiles of thyroid carcinomas and to explore their expression in FVPTC arising in follicular adenomatoid lesion so as to validate the altered expression of genes in the pathobiology of PTC development.

References:

1. Davies L, Welch G. Increasing incidence of thyroid cancer in the United States, 1973-2002. JAMA 295(18): 2164-2167, 2006.

2. Nikiforov YE.Molecular analysis of thyroid tumors.Mod Pathol 24: Suppl 2:S34-43, 2001.

New insights on bone biology from exome sequencing of osteopetrotic patients with atypical presentations

Eleonora Palagano1,2, Alessandra Pangrazio1,2,Dario Strina1,2, Alessandro Puddu3,4, Manuela Oppo3, Maria Valentini3, Gianmauro Cuccuru3, Paolo Uva3, Andrea Angius3,4, Paolo Vezzoni1,2, Anna Villa1,2, Cristina Sobacchi1,2

1CNR-IRGB, Milan Unit, Italy; 2Istituto Clinico Humanitas IRCCS, Rozzano, Italy;
3CRS4 , Parco Scientifico e Tecnologico POLARIS, Pula, Italy;
4CNR-IRGB, Cittadella Universitaria di Monserrato, Cagliari, Italy

Autosomal Recessive Osteopetrosis (ARO) is a rare genetic bone disease presenting early in life with extreme sclerosis of the skeleton, reduction of bone marrow spaces, hepatosplenomegaly, cranial nerves compression and severe growth failure. ARO is often lethal within the first decade of life because of anaemia and secondary infections; at present, the only therapeutic option is Hematopoietic Stem Cell Transplantation, which should be performed as soon as possible in order to obtain a major benefit by avoiding the establishment of irreversible defects. Human ARO is genetically heterogeneous; mutations in 7 different genes have been described (*TCIRG1*, *CLCN7*, *OSTM1*, *PLEKHM1*, *SNX10*, *RANKL* and *RANK*), with crucial roles in either osteoclast function or differentiation. Unfortunately, in some cases delays in diagnosis occur, due to ARO rareness and to the presence of complex phenotypes which may be misinterpreted. In this context, the identification of the specific molecular defect allows to unequivocally draw the clinical diagnosis and establish the appropriate treatment.

The recent development and broad use of next generation sequencing techniques has importantly impacted on disease gene identification in many inherited pathologies. In particular, we have recently undertaken exome sequencing in osteopetrotic patients lacking a genetic classification. Here we report the results obtained in 3 unrelated families using this approach. In family 1, presenting 2 affected siblings initially diagnosed with “intermediate osteopetrosis”, we identified a homozygous mutation in the *CTSK* gene. Prompted by this finding, we tested by Sanger sequencing 25 additional patients and found *CTSK* mutations in 4 of them, whose clinical and radiographic features were retrospectively found to be compatible with, but

not typical for, Pycnodysostosis. On this basis, we recommend that *CTSK* gene be included in the molecular diagnosis of intermediate forms of human ARO and, more in general, of high-density bone conditions.

In family 2, exome sequencing revealed in the patient the presence of a novel homozygous mutation in the Carbonic Anhydrase II (*CA2*) gene, which was an unexpected finding, since both MRI and CT scan did not show in the child the cerebral calcifications typical for this subset of ARO. While further clinical investigations are on-going in the patient, our data would suggest a certain degree of variability of presentation also in CA2-dependent ARO.

In family 3, exome sequencing highlighted the presence of a single nucleotide variant in the low-density lipoprotein receptor-related protein 5 (*LRP5*) gene, predicted to cause a non-conservative amino acid change, at the homozygous state in the 2 affected siblings. Defects in this gene are known to cause two opposite bone phenotypes: recessive loss-of-function mutations lead to osteoporosis pseudoglioma (OPPG) syndrome, which is a juvenile-onset autosomal recessive disease of low bone mass; on the contrary, dominant gain-of-function mutations lead to high bone mass (HBM) phenotypes. In addition, common polymorphic variants in LRP5 have been associated with bone mineral density in several reports. Interestingly, the nucleotide change identified in family 3 lies in a region of the protein that is not reportedly associated with increased bone density; furthermore it appears to be recessive in nature. On this basis, further investigation is on-going in order to verify the causative role of this mutation and, in case, unravel the exact molecular mechanisms leading to the disease.

Overall, we underline the difficulties of differential diagnosis in some patients whose clinical appearance does not fit the classical malignant or benign picture, and confirm the role of exome sequencing in the molecular classification of genetically heterogeneous diseases.

A case report of a boy with unusual presentation of cleidocranial dysplasia - whole exome sequencing ends 10 years of diagnostic odyssey or does it really?

Sander Pajusalu[1], Margit Nõukas[3,4], Tiina Rägo[5], Andres Metspalu[3,4], Katrin Õunap[1,2]

[1]Department of Genetics, United Laboratories, Tartu University Hospital, Tartu, Estonia
[2]Department of Pediatrics, University of Tartu, Tartu, Estonia
[3]Estonian Genome Center, University of Tartu, Tartu, Estonia
[4]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
[5]Children's Clinic, Tartu University Hospital, Tartu, Estonia

At the beginning of the genomic era in clinic genetics whole exome sequencing (WES) is used enthusiastically to resolve long diagnostic challenges – cases that have stayed unresolved despite exhaustive genetic, metabolic and other testing. Still, the interpretation of whole exome results in those cases may remain controversial as there is a possibility to find genetic diagnosis which does not explain the whole clinical phenotype of the patient.

Here we present a case of a 10-year-old boy with complex phenotype who was first seen by clinical geneticist as a neonate and then regularly since today. He presented antenatally with the ultrasound findings of short extremities and polyhydramnion. After birth from Caesarean section widely opened sutures and

large fontanels were documented as well as many dysmorphic features. Since 3 weeks of age he developed jaundice with hepatosplenomegaly. The liver biopsy confirmed the paucity of extrahepatic bile ducts; additionally liver fibrosis and polysplenia was found. At the age of 1.5 months the hepatoportoenterostomy was done which gave positive therapeutic effect. Afterwards he has had features of psychomotor developmental delay, hepatobiliary symptoms, low weight and small stature, mild failure to thrive, many facial dysmorphic features, pupil asymmetry, supranumerary teeth with malocclusion, relatively short extremities and hip dysplasia, hyperelastic skin, joint hypermobility, crab-like thorax, short fingers with broad fingertips.

During the years many cytogenetic (karyotype, FISH for *JAG1* gene, chromosomal microarray), molecular genetic (SEC23A gene sequencing), metabolic and even electron microscopy studies of endoplasmatic reticulum did not reveal the cause of the disorder. The last unconfirmed diagnostic hypothesis was cranio-lenticulo-sutural dysplasia. After consulting with many experts, WES was suggested as a next step in diagnostic pathway.

Whole exome sequencing revealed mutation in RUNX2 gene (c.674G>A, p.Arg225Gln) known to cause cleidocranial dysplasia (CCD). CCD was previously considered as a differential diagnosis, but was excluded on clinical ground because CCD does not explain biliary atresia, hepatosplenomegaly, polysplenia and connective tissue symptoms.

As WES revealed a mutation only partially explaining the complex phenotype, the question remains whether the patient has unusual previously undescribed form of CCD or maybe he has two different conditions at once. We can always turn back to WES data to search for "second" mutation, but is it worth it?

Comprehensive massive parallel DNA sequencing strategy for the genetic diagnosis of the Neuro-cardio-facio-cutaneous syndromes

Ana Justino[1,2], Patrícia Dias[3], Maria João Pina[1], Sónia Sousa[1], Luís Cirnes[1], Ana Berta Sousa[3], José Carlos Machado[1,4] and José Luis Costa[1,4]

[1] Institute of Molecular Pathology and Immunology of the University of Porto, Portugal

[2] Abel Salazar Institute for the Biomedical Sciences of the University of Porto, Portugal

[3] Department of Genetics, Hospital de Santa Maria de Lisboa, Portugal

[4] Faculty of Medicine of the University of Porto, Portugal

Variants in 11 genes of the RAS/MAPK signaling pathway have been causally linked to the Neuro-cardio-facio-cutaneous syndromes group (NCFCS). Recently, *A2ML1* and *RIT1* were also associated with these syndromes. Due to the genetic and clinical heterogeneity of NCFCS it is challenging to define strategies for their molecular diagnosis. The aim of this study was to develop and validate a massive parallel sequencing (MPS) based strategy for the molecular diagnosis of NCFCS. A multiplex PCR-based strategy for the

enrichment of the 13 genes and a variant prioritization pipeline was established. Two sets of genomic DNA samples were studied using the Ion PGM System: (a) training set (n=15) to optimize the strategy; (b) validation set (n=20) to validate and evaluate the power of the new methodology. Sanger sequencing was performed to confirm all variants and low covered regions. All variants identified by Sanger sequencing were detected with our MPS approach. The methodology resulted in an experimental approach with a specificity of 99.0% and a maximum analytical sensitivity ≥98.2% with a confidence of 99%. Importantly, two patients (out of 20) harbored described disease causing variants in genes that are not routinely tested (*RIT1* and *SHOC2*). The addition of less frequently altered genes incremented in ≈10% the diagnostic yield of the strategy currently used. The presented workflow provides a comprehensive genetic screening strategy for patients with NCFCS in a fast and cost-efficient manner. This approach demonstrates the potential of a combined MPS-Sanger sequencing based strategy as an effective diagnostic tool for heterogeneous diseases.

Missense Mutation in FNDC3A causes Split Hand/Foot Malformation

Sinje Geuer1,2, Sandra C. Dölken1, Aleksander Jamsheer3,4, Peter Krawitz1,2, Jochen Hecht5, Silke Lohan2, Eva Klopocki1,6, Peter N. Robinson1,2, Stefan Mundlos1,2

1)Institute for Medical and Human Genetics, Charité Universitätsmedizin Berlin, Germany 2)Max Planck Institute for Molecular Genetics, Berlin, Germany 3)Department of Medical Genetics, Poznan University of Medical Sciences,Poznan, Poland 4)NZOZ Center for Medical Genetics GENESIS, Poznan, Poland 5)Berlin-Brandenburg Center for Regenerative Therapies (BCRT), Charité – Universitätsmedizin Berlin, Berlin, Germany 6)Institute for Human Genetics, Universität Würzburg, Germany

Split hand/foot malformation (SHFM) is a congenital limb malformation characterized by truncation or loss of central rays of hands and/or feet. It occurs in syndromic- and non-syndromic forms; dominant, recessive and in rare cases x-linked inheritance have been reported. So far, six different loci (SHFM1-6) have been associated with SHFM: point mutations in *TP63* and *WNT10B*, copy number changes in 10q24, 17p13.3 and 7q21 as well as translocations in 7q21; but in a large number of cases the underlying cause remains unresolved.

We studied a large consanguineous family from Syria with four individuals affected by non-syndromic SHFM. Mutations in the known loci had been excluded previously. We identified a homozygous mutation in the *FNDC3A* gene (*Fibronectin Domain Containing Protein 3A*) by whole exome and subsequent confirmation by Sanger sequencing. All affected family members share the homozygous missense mutation in *FNDC3A* while the nine analyzed non-affected members carry either only a heterozygous mutation or wildtype.

FNDC3A is thought to be important for cell-cell adhesion (Obholz et al. 2006) which is critical for the development of different tissues. For the limb, the adhesiveness of mesenchymal cells undergoes spatiotemporal changes during cartilage formation (Wada et al., 2011) and alterations in cell adhesion have

been shown to lead to limb malformations such as distal truncations (Yamaguchi et al., 1999). The mutation is located in a highly conserved region of *FNDC3A*, in the last of nine fibronectin domains. We could show by whole mount *RNA in situ* staining mouse embryos that *FNDC3A* is expressed in the Apical Ectodermal Ridge of the developing limb bud. This region is known to be essential for outgrowth and patterning of the limb. In patient fibroblast FNDC3A is downregulated, while the Knockout in Zebrafish leads to a truncation of the caudal fin, mimicking SHFM.

In summary, we suggest *FNDC3A* as a candidate gene for autosomal recessive SHFM.

A Mouse Model to reconstruct Craniosynostosis and Syndactyly associated with CNVs within the Regulatory Landscape of IHH

Anja Will

Max-Planck Insitute of Molecular Genetics, Berlin Germany

Indian hedgehog (IHH) is expressed in the growth plate and plays an essential role in the development of bone and inner organs. Mutations of *IHH* are associated in humans with brachydactyly. Besides point mutations, copy number variations (CNVs) represent structural variations in the human genome. Deletions or duplications of genes and of their regulatory landscape can strongly affect gene expression levels and have been shown to be associated with diseases. The expression of developmental genes like *IHH* is known to rely on regulatory elements in the non-coding region of the gene locus. Using array comparative genomic hybridization our group showed that CNVs i.e. duplications within the 5' non-coding region of the *IHH* locus are associated with syndactyly and craniosynostosis Philadelphia type. Syndactyly is characterized by the fusion of digits, either by ossification or by soft tissue (Bosse *et al*, 2000). Craniosynostosis is characterized by the premature fusion of the fibrous sutures of an infant skull and thus, restricts the growth of skull and brain (Slater *et al*, 2008; Gault *et al*, 1992). Within the duplications we could identify three regions (conserved elements CE1-3), which are highly conserved in human, mouse and chicken. We assume that the duplication of these elements leads to a mis-expression of *IHH* and therefore strongly influences the development of the digits and the skull. To study the complexity of genomic aberrations at the *IHH* locus and to completely understand the resulting phenotype we want to design a mouse model which enables us to duplicate one, two or all three CEs. We want to generate three mouse alleles that contain an insertion with a loxP site at the breakpoints of the duplications. By Cre/loxP-mediated recombination we want to recapitulate the human phenotype in mice. Furthermore, the LacZ-reporter of the insert allows us to identify regulatory domains. In addition, we want to remobilize the insertion cassette using a Sleeping Beauty (SB)-transposon approach to study the importance of potential regulatory elements within the *IHH* locus.

FACULTY Who is Who

Family Name	Name	Affiliation	Work e-mail
Ameur	Adam	Science for Life Laboratory, National Genomics Infrastructure, Uppsala University, Sweden	adam.ameur@igp.uu.se
Bonnefond	Amelie	CNRS UMR8199, Lille Pasteur Institute, Lille II University, Lille, France	amelie@good.ibl.fr
Boustred	Christopher	North East Thames Regional Genetics Laboratory, Great Ormond Street Hospital NHS Foundation Trust, London	cboustred@gmail.com
Calabrese	Francesco	CEINGE (Centro Ingegneria Genetica) Advanced Biotechnologies, Naples, Italy and Dip. Bioscienze, Biotecnologie e Biofarmaceutica, Univ. Of Bari	franccalabrese@libero.it
De Bakker	Paul	Departments of Medical Genetics and of Epidemiology University Medical Center Utrecht The Netherlands	P.I.W.deBakker-2@umcutrecht.nl
Gilissen	Christian	Department of Human Genetics - Nijmegen Centre for Molecular Life Sciences- Radboud University Nijmegen Medical Centre, the Netherlands	c.gilissen@antrg.umcn.nl
Matthijs	Gert	Center For Human Genetics, University of Leuven, Belgium	gert.matthijs@med.kuleuven.be
Pippucci	Tommaso	University of Bologna, Italy	tommaso.pippucci@unibo.it
Rauch	Anita	Anita Rauch, University of Zurich, Institute of Medical Genetics, Schlieren-Zurich, Switzerland	anita.rauch@medgen.uzh.ch

Robinson	Peter	Institut für Medizinische Genetik und Humangenetik Charité - Universitätsmedizin Berlin Germany	peter.robinson@charite.de
Romeo	Giovanni	University of Bologna, Italy	romeo@eurogene.org
Sidore	Carlo	CNR, Institute of Genetics Research, Sassari, Italy	scarlino81@gmail.com
Tomlinson	Ian	Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN	iant@well.ox.ac.uk
Veltman	Joris	Department of Human Genetics- Nijmegen Centre for Molecular Life Sciences Radboud University Nijmegen Medical Centre	J.Veltman@antrg.umcn.nl

Who is Who: Participants

NAME	FAMILY NAME	AFFILIATION	NATIONALITY	CITTA'	E-MAIL
MARYAM	AL NABHANI	Sultan Qaboos University Hospital	OMAN	MUSCAT	alnabhani_mr@hotmail.com
ANWAR	BABAN	Bambino Gesù Pediatric Hospital	ITALY	ROME	anwar.baban@opbg.net
GIULIANO	BOLONDI	CSS Mendel	ITALY	ROME	g.bolondi@css-mendel.it
MARTINA	CALORE	Univ. Of Padova	ITALY	PADOVA	martina.calore@unipd.it
CHIARA MAURA	CINISELLI	IRCCS Foundation	ITALY	MILAN	chiara.ciniselli@istitutotumori.mi.it
FRANCESCA	CORTINI	Maggiore Policlinico Hospital	ITALY	MILAN	francescacortini@gmail.com
MIRELLA IRENE STELLA	CRAPANZANO	Policlinico Universitario Messina	ITALY	MESSINA	mirella.crapanzano@hotmail.it
SONIA ALEXANDRA	CUSTODIO	Hospital de Santa Maria	PORTUGAL	CARCAVELOS	sonia.alex.custodio@gmail.com
GORAN	CUTURILO	University Children's Hospital	SERBIA	BELGRADE	udkgenetika@udk.bg.ac.rs
MARYAM	DANESHPOUR	Research Institute of Endocrine Research	IRAN	TEHERAN	daneshpour1388@gmail.com
NADA	DANIAL FARRAN	Ha'Emek Medical Center	ISRAEL	NAZARETH ILLIT	nada_fa@clalit.org.il
MARIACRISTINA	DI MARINO	Mario Negri Institute	ITALY	NAPOLI	mariacristina.dimarino@marionegri.it
ANA CAROLINA	DOS SANTOS FONSECA	Univ. Of San Paulo / Univ. Of Copenhagen	DENMARK	COPENHAGEN	ana.fonseca@usp.br
JAKUB	FICHNA	Mossakowski Medical Research Centre	POLAND	WARSAW	jfichna@imdik.pan.pl
SINJE	GEUER	Max-Planck Insitute of Molecular Genetics	GERMANY	BERLIN	geuer@molgen.mpg.de
LONNEKE	HAER-WIGMAN	Radboudumc	NETHERLANDS	NIJMEGEN	Lonneke.haer-wigman@radboudumc.nl
GEIR ASMUND	HANSEN	Univ. Of Nothern Norway	NORWAY	TROMSO	geir.asmund.hansen@unn.no
DOROTA	HOFFMAN ZACHARSKA	Institute of Mother and Child	POLAND	WARSAW	dhoffman@wp.pl
ALLAN	HOJLAND	Aalborg University Hospital	DENMARK	AALBORG	a.hoejland@rn.dk
ADRIANO	JIMENEZ ESCRIG	FIBIO-HRC	SPAIN	MADRID	adriano.jimenez@hrc.es
ALED	JONES	Guys St. Thomas's NHS Trust	UK	FLITWICK	aledj2@hotmail.com
CHRISTOFFER	JONSRUD	University Hospital of North Norway	NORWAY	TROMSO	christoffer.jonsrud@unn.no
ANA	JUSTINO	IPATIMUP	PORTUGAL	PORTO	ajustino@ipatimup.pt
RABAB	KHAIRAT	National Research Centre	EGYPT	CAIRO	rababncr2001@hotmail.com
KAMAL	KHAZANEHDARI	MBG Laboratory	EMIRATI ARABI	DUBAI	kamalk@mbg.ae
GASPER	KLANCAR	University Medical Center Ljubljana	SLOVENIA	LJUBLJANA	gasper.klancar@kclj.si
JERNEY	KOVAC	University Medical Center Ljubljana	SLOVENIA	LJUBLJANA	alenka@airpass.eu

MONIKA	LAMBA SAINI	Cliniques Universitaires Saint-Luc	BELGIUM	BRUXELLS	monikalamba@gmail.com
ITALIA	LODDO	University of Messina	ITALY	MESSINA	italia.loddo@gmail.com
IHAB	LOLAS	Aalborg University Hospital	DENMARK	AALBORG	i.lolas@rn.dk
LAURA	LORES DE MOTTA	Radboud University Medical Center	NETHERLANDS	NIJMEGEN	laura.loresdemotta@radboudumc.nl
CELIA	NOGUEIRA	National Institute of Health	PORTUGAL	PORTO	celianogueira55@hotmail.com
BJORN	NYGARD	Dept. Of Medical Genetics of North Norway	NORWAY	TROMSO	bjorn.nygard@unn.no
FRANCESCO	ORSINI	Univ. Of Milan	ITALY	MILAN	francesco.orsini1@unimi.it
LUCA	PACINI	Univ. Of Rome La Sapienza	ITALY	LATINA	luca.pacini@uniroma1.it
SANDER	PAJUSALU	Tartu University Hospital	ESTONIA	TARTU	sanderpajusalu@gmail.com
ELEONORA	PALAGANO	Humanitas	ITALY	FLORENCE	eleonora.palagano@gmail.com
EVANGELIA	PANAGIOTOU	University of Leeds	NETHERLANDS	LEEDS	e.panagiotou@leeds.ac.uk
PIA	POHIOLA	TYKS-SAPA	FINLAND	TURKU	vellai@utu.fi
GIULIA	POLONI	Univ of Padova	ITALY	PADOVA	giuliapoloni7@gmail.com
LARS	RETTERSTOL	Oslo University Hospital	NORWAY	OSLO	lars.retterstol@ous-hf.no
SETAREH	SAFAVI	Lunds University	SWEDEN	LUND	setareh.safavi@med.lu.se
RICCARDO	SANGERMANO	University Medical Center St. Radboud	NETHERLANDS	NIJMEGEN	Riccardo.Sangermano@radboudumc.nl
PATRICK	SCOTT	SQUH	OMAN	MUSCAT	patrick_scott32@yahoo.com
NAVA	SHAUL LOTAN	Hadassah Hebrew University Medical Center	ISRAEL	JERUSALEM	navashaul34@gmail.com
ROSARIO	SILVEIRA SANTOS	Hospital Santa Maria	PORTUGAL	LISBON	rosariosil1@gmail.com
MARGJE	SINNEMA	Dept. Of Clinical Genetics MUMC	NETHERLANDS	MAASTRICHT	margje.sinnema@mumc.nl
NADIA	SKAULI	University of Oslo	NORWAY	OSLO	nadia.skauli@mn.uio.no
BEATE	SKINNINGSRUD	University of Oslo	NORWAY	OSLO	beate.skinningsrud@medisin.uio.no
SONIA	SOUSA	IPATIMUP diagnostics	PORTUGAL	PORTO	ssousa@ipatimup.pt
MARKETA	TESAROVA	Prague University Hospital	CZECK REPUBLIC	PRAGUE	marketa.tesarova@lf1.cuni.cz
MINNA	TOIVONEN	University of Turku	FINLAND	TURKU	minna.toivonen@utu.fi
KATARINA	TREBUSAK PODKRAJSEK	University Medical Center Ljubljana	SLOVENIA	LJUBLJANA	alenka@airpass.eu
JAN	VEJVALKA	Datim, spol. S.r..o.	CZECK REPUBLIC	PRAGUE	jan.vejvalka@lfmotol.cuni.cz
DIDEBERG	VINCIANE	Genetique Humaine - CHU Liege	BELGIUM	LIEGE	vinciane.dideberg@chu.ulg.ac.be
MONIKA	WEISZ HUBSHMAN	Schneider	ISRAEL	PETAH TIKVA	whmonika@gmail.com
ANJA	WILL	Max-Planck Insitute of Molecular Genetics	GERMANY	BERLIN	will_a@molgen.mpg.de
DANIELA	ZAHORAKOVA	Charles University Hospital	CZECK REPUBLIC	PRAGUE	daniela.zahorakova@vfn.cz
MARKUS	ZWEIER	University of Zurich	SWITZERLAND	SCHLIEREN	mazweier@gmx.de