

MEETING REPORT

Next-generation sequencing for identifying new genes in rare genetic diseases: many challenges and a pinch of luck

Amélie Bonnefond^{1,2,3*} and Philippe Froguel^{1,2,3,4,5*}

Abstract

A report on the European Society of Human Genetics conference, held in Paris, France, June 8-11, 2013.

Keywords: Bioinformatics; genome sequencing; next-generation sequencing; prioritization of variants; rare disorder; syndrome; whole-exome sequencing

The European Society of Human Genetics is a non-profit organization aiming to promote research in basic and applied human and medical genetics. This year, there were almost 3,000 attendees at the annual meeting, which reflected state of the art technologies, innovations, novel strategies, counseling and applications in human and medical genetics. As expected, a plethora of genetic diseases was tackled: from very rare disorders (such as developmental syndromes, neuromuscular disorders, metabolic and mitochondrial disorders, dysmorphic syndromes, monogenic forms of epilepsies, retinal dystrophies), to more common diseases (such as cancer, autism, intellectual disability).

In this report, we focus on the several sessions, symposia and workshops that discussed next-generation sequencing (NGS) technologies and their contribution to the identification of new genetic etiologies of rare disorders. The current challenges of this technology are also discussed.

Success in finding new genetic etiologies in rare diseases through NGS

During the workshop on NGS in clinical practice, clinical geneticists from three European laboratories illustrated their diagnostic NGS approaches and results obtained so far. They mostly discussed generic whole-exome sequencing (WES) in severe syndromic cases, and two main filtering strategies for identifying the causal variant(s)

were used in those cases: strategies focusing on *de novo* mutations (not observed in the parents) and strategies on homozygous mutations (in patients from consanguineous families), which have not been reported in public databases and are predicted to be damaging. Using this approach, Lissenka Vissers (Radboud University Nijmegen Medical Centre, The Netherlands) and collaborators identified a *de novo* mutation in the *ARID1B* gene, in a patient presenting with eczema, short stature, delayed development and intellectual disability. It is known that mutations in this gene cause Coffis-Siris syndrome, a disease with an autosomal recessive inheritance. Furthermore, the same group identified a *de novo* mutation in the *PDHAI* gene, in a patient presenting with microcephaly, feeding problems, severe developmental delay (non-speech), behavior problems, intellectual disability, short stature, hirsutism and paraplegia. It is of note that before WES, the *PDHAI* gene had been erroneously reported negative in the patient via Sanger sequencing.

Anita Rauch (Zürich University, Switzerland) and collaborators discussed a case presenting with Fraser-syndrome-like clinical features (including kidney and ureter agenesis). Four genes (including 100 exons) can be mutated in this syndrome. The assessment of these genes would have been too expensive and taken too much time using Sanger sequencing. Therefore, they used WES and found, after filters, a truncated variant and a substitution, both located in one of the four known genes: *FRAS1*. They discussed another patient with intellectual disability, severe short stature and severe microcephaly, and born from consanguineous parents. They found a homozygous mutation of interest in the *CRNPJ* gene that is known to be mutated in Seckel syndrome. However, the overlap of phenotypes was incomplete. Interestingly, they found another homozygous mutation of interest in the *SACS* gene and suggested a compound phenotype that would explain all disorders of the patient.

Rauch and collaborators also investigated a third case presenting with macrosomia, severe hypotonia, hyperactivity, autistic features and postaxial hexadactyly. They found a *de novo* frameshift in *OFD1* that causes oral-

*Correspondence: amelie.bonnefond@good.ibl.fr; philippe.froguel@good.ibl.fr

¹CNRS-UMR8199, Lille Pasteur Institute, Lille 59000, France

Full list of author information is available at the end of the article

facial-digital syndrome type I when mutated. However, the phenotypes were not overlapping. They subsequently looked for mosaic mutations in a candidate gene (that had just been published): *PIK3CA*. They successfully identified a mosaic substitution in this gene that was present in 16% of the saliva cells and in 7% of the blood cells.

In a concurrent session, Jean-Baptiste Rivière and collaborators (Laboratoire de Génétique Moléculaire, CHU, France) presented a study in which they performed WES in two patients (and their unaffected parents) with short stature, hyperextensibility of joints and/or inguinal hernia, ocular depression, Rieger anomaly and teething delay (SHORT) syndrome. By focusing on *de novo* mutations of interest, they found two different substitutions in *PIK3R1*. Subsequently, via Sanger sequencing, they found seven additional carriers of *PIK3R1* mutations who presented with SHORT syndrome. Furthermore, they demonstrated the functionality of these mutations using fibroblasts derived from individuals with *PIK3R1* mutations. Their findings highlighted the critical role of *PIK3R1* in insulin action, and normal growth and development.

In the same session, Sian Ellard and collaborators (University of Exeter Medical School, UK) used a combination of genome sequencing and homozygosity mapping analysis in consanguineous pedigrees affected with neonatal diabetes linked to pancreatic agenesis. They found a shared run of homozygosity encompassing the *PTF1A* gene known to cause pancreatic agenesis when mutated. No mutations were found in the *PTF1A* gene. However, through genome sequencing in two probands, they identified a substitution located in a highly conserved 400 bp region. Through functional analyses, they showed that this region was actually an enhancer of *PTF1A*. Four additional mutations and one indel spanning the same region were identified in ten affected families. This study shows the potentially key role of non-coding regions in monogenic disorders.

Other NGS success stories in rare disorders were presented at the meeting. However, it is noteworthy that the NGS results were rather disappointing in common diseases and in general populations. A concurrent session, including six oral presentations, was devoted to this topic, but no strong data were presented. However, most of the presenters tackled the problem of statistical power.

Challenges in NGS

In the educational session named 'NGS in the clinic', Christian Gilissen (Radboud University Nijmegen Medical Centre, The Netherlands) tackled several challenges in NGS and discussed with the audience what can go wrong in NGS analyses. It was noted that sample quality can have detrimental effects in NGS analyses, affecting the transition:transversion ratio, overlap with public databases, increases in private variants, as well as negative

effects on fastQC data. Gilissen also discussed how not sequencing variants can be due to a low coverage, which is determined by the number of reads for a sample, the number of reads that can be aligned to the genome and the number of reads mapping to the exome (enrichment). Of note, some regions do not enrich well, but in good samples 90% of reads can efficiently map to the target. However, there can be specific enrichment problems due to, for instance, GC-rich regions. False-positive variants can also be due to high amounts of sequence errors (even when there is high coverage), a homopolymer stretch, a PCR amplification artifact and strand bias.

Gilissen and Yves Moreau (University of Leuven, Leuven, Belgium) discussed the big challenge in prioritizing variants. After standard filtering against variants that are not present in healthy populations, many potential candidate mutations are often retained and predictive methods to prioritize variants for further validation are needed. The current computational methods take into account biochemical, evolutionary and structural properties of mutations to assess their functional effect. However, most of these methods overpredict the pathogenicity of the mutations. Moreau and collaborators suggested a genomic data fusion methodology (Extasy) that integrates multiple strategies to detect deleteriousness of mutations and prioritizes them in a phenotype-specific manner using diverse genomic sources such as Gene Ontology (GO), Kyoto Encyclopedia of Genes and genomes (KEGG) and InterPro.

Conclusion

To conclude, although many NGS success stories were reported during this European Society of Human Genetics meeting, these results were only found in monogenic diseases, especially recessive or *de novo* dominant disorders. NGS data in polygenic diseases have been very disappointing to date, and so it is clear that we still have many challenges to overcome in NGS.

Abbreviations

bp, base pair; NGS, next-generation sequencing; SHORT syndrome, short stature, hyperextensibility of joints and/or inguinal hernia, ocular depression, Rieger anomaly and teething delay syndrome; WES, whole-exome sequencing.

Competing interests

The authors declare that they have no competing interests.

Author details

¹CNRS-UMR8199, Lille Pasteur Institute, Lille 59000, France. ²Lille II University, Lille 59000, France. ³European Genomic Institute for Diabetes (EGID), Lille 59000, France. ⁴Department of Genomics of Common Disease, School of Public Health, Imperial College London, Hammersmith Hospital, London W12 0NN, UK. ⁵Qatar Biomedical Research Institute (QBRI), Qatar Foundation, Doha, Qatar.

Published: 29 July 2013

doi:10.1186/gb-2013-14-7-309

Cite this article as: Bonnefond A, Froguel P: Next-generation sequencing for identifying new genes in rare genetic diseases: many challenges and a pinch of luck. *Genome Biology* 2013, **14**:309.