Rare Diseases and the NHS 100,000 Genomes Project

Willem H Ouwehand MD PhD FMedSci
Professor of Experimental Haematology
Director NIHR BioResource – Rare Diseases
The NHS 100,000 Genome Project

- Birth/Year: ~800,000 (80%)
- Inherited Diseases: 40,000

NIHR BioResource
Rare Diseases
The NHS 100,000 Genome Project

NIHR - BioResource
- Ethics
- Facilitating Enrolment
- Phenotype Data Capture

NIHR - National Biosample Centre

Sample

NHS

Inherited Diseases
40,000

NIHR BioResource Rare Diseases

Birth/Year
~800,000

20%
The NHS 100,000 Genome Project

NIHR - BioResource
- Ethics
- Facilitating Enrolment
- Phenotype Data Capture

NIHR - National Biosample Centre

NIHR BioResource
Rare Diseases

NIHR BioResource
Inherited Diseases 40,000

NIHR BioResource
Birth/Year ~800,000

NIHR BioResource
Sample

Small Data Report

illumina®
Sequence Provider

Academia
- Maths
- Computer Science
- Physics

EBI
Annotate the genome

Genomics England Ltd
To make genomic medicine happen
GECIP = Genomics England Clinical Interpretation Partnership

Big Data Flow
The NIHR National Biosample Centre
a highly automated and scalable sample handling, testing and archiving facility
Whole Genome Sequencing (WGS) results

mean minimum percentage of genome covered

~2% covered <15x

1,789 samples
Whole Genome (WGS) vs Whole Exome Sequencing (WES)

first 10 Mb of chromosome 10
Whole Genome (WGS) vs Whole Exome Sequencing (WES)

first 10 Mb of chromosome 10
## WES vs WGS

*comparison of results in Rare Diseases pilot*

<table>
<thead>
<tr>
<th>Genotyping Method</th>
<th>Number of Samples</th>
<th>All variants (in M)</th>
<th>Rare coding variants &lt;1:1000 *</th>
<th>% increase in sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>WES</td>
<td>737</td>
<td>3.3</td>
<td>129,394</td>
<td>100</td>
</tr>
<tr>
<td>WGS</td>
<td>736</td>
<td>45.3</td>
<td>176,592</td>
<td>136</td>
</tr>
</tbody>
</table>

* Calculated over 64 Mb of coding space
Platelets
- 150 – 400 x $10^9$/L
- Circulate as discoid cells

1 µm
Imbalance between damage and repair is the Number 1 killer in Western society
Three gene discovery stories
with three new messages, one about biology and two about genetic architecture

~200,000 base pairs deleted (10 genes)
A mutation from both parents

Grey Platelets &
defective alpha granules
*NBEAL2*
Albers, Nat Genet 2011

- Known since 1970
- ~50 cases worldwide
- Moderate bleeding
- Large and low number of platelets
- Absence of α-granules in platelets
Faithful phenocopy

Scarring of the blood stem cell niche and a pro-inflammatory state

Guerrero et al. Blood 2014
Faithful phenocopy

Scarring of the blood stem cell niche and a pro-inflammatory state

No metastasis of melanoma cells to the lungs

Lack of α-granules has the most profound effect ever observed in this model
A deletion from one parent & not-so-rare regulatory SNP from the other parent

VEL group & new genetic mechanism
**SMIM1**
Cvejic, *Nat Genet* 2013

TAR & new genetic mechanism
**RBM8A**
Albers, *Nat Genet* 2012
Albers, *Curr Op Gen* 2013
Chromatin marks and epigenetic regions drive differentiation programs and confer cellular identity and functional phenotypes.
IMMUNOLOGY

Metabolic shift may train immune cells

BLUEPRINT project studies epigenetics of various blood cells

RESEARCH ARTICLE

mTOR- and HIF-1α-mediated aerobic glycolysis as metabolic basis for trained immunity

Shih-Chin Cheng,1 Jessica Quintin,2 Robert A. Cramer,2 Kelly M. Shepardson,2 Sadia Saeed,3 Vinod Kumar,4 Evangelos J. Giamarellos-Bourboulis,5 Joost H. A. Martens,3 Nagesha Appukutla Rao,6 Ali Aghajanirefah,7 Ganesh R. Manjeri,4 Yang Li,4 Daniela C. Iftin,4 Rob J. W. Arts,1 Brian M. J. W. van der Meer,6 Peter M. T. Deen,7 Colin Logie,5 Luke A. O'Neill,6 Peter Willems,6 Frank L. van de Veerdonk,1 Jos W. M. van der Meer,1 Aylwin Ng,9,30 Leo A. B. Joosten,4 Circa Wijmenga,4 Hendrik G. Stunnenberg,4 Ramnik J. Xavier,5,9,10 Mihai G. Netea11

IMMUNOGENETICS

Transcriptional diversity during lineage commitment of human blood progenitors

Lu Chen,1,2,8 Myro Kostadima,2,4,39 Joost H. A. Martens,3,4 Giovanni Camu,2,3 Sara P. Garcia,2,3 Ernest Turro,2,3 Kate Downes,2,3 Iain Macaulay,6 Ewa Bielezyk-Maczynska,2,2 Sophia Coe,2,3,4 Samantha Farrow,2,3 Pawan Poudel,2,3 Frances Burden,2,3 Sjoert B. G. Jansen,2,3 William J. Astle,2,3,7 Antony Attwood,2,3 Tadibir Bariana,8,9 Bernard de Bono,10,11 Alessandra Breschi,12 John C. Chambers,13,14 BRIDGE Consortium,1 Fizzah A. Choudry,2,3,4 Laura Clarke,4 Paul Coupland,1 Martijn van der Ent,5 Wendy N. Erber,15 Joop H. Jansen,16 Rémi Fayvet,17 Matthew E. Fenech,18 Nicola Foad,2,3,4,4 Kathleen Freson,19 Chris van Geet,19 Keith Gomez,4 Roderic Guigo,12 Daniel Hampp,1,2,3 Anne M. Kelly,2,3,20 Hindrik H. D. Kerstens,5 Jaspal S. Koenderman,15,14 Michael Laffan,31 Claire Lenthall,21 Charlotte Labalette,2,3 Tiphaine Martin,2,3,22 Stuart Meacham,9,23 Andrew Mumford,25 Sylvia Nürnberg,18,24 Emilio Palumbo,12 Bert A. van der Reijden,16 David Richardson,1 Stephen J. Sambur,16,44,45 Greg Slodkowski,4 Asif U. Tamuri,5 Louella Vasquez,11 Katrin Voss,9,36 Stephen Watt,7 Sarah Westbury,25 Paul Flicek,4,3 Remco Loos,4 Nick Goldman,4 Paul Bertone,4,27,28 Randy J. Read,29 Sylvia Richardson,7 Ana Cvejic,3,1 Nicola Soranzo,1,5,1 Willem H. Ouwehand,5,2 Hendrik G. Stunnenberg,5,2,3 Mattia Frontini,2,3,4,11 Augusto Rondon2,3,4,11

RESEARCH ARTICLE

Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity

Sadia Saeed,1 Jessica Quintin,2* Hindrik H. D. Kerstens,1* Nagesha A. Rao,1* Ali Aghajanirefah,1* Filomena Matarese,1 Shih-Chin Cheng,2 Jacqueline Ratter,2 Kim Berentsen,1 Martijn A. van der Ent,1 Niko Jaffe1, Eva M. Janssen-Megen,1 Menno Ter Huurre,1 Amel Mandoli,3 Tom van Schaijk,1 Aylwin Ng,9,34 Frances Burden,5,6 Kate Downes,2,39 Mattia Frontini,2,3 Vinod Kumar,4 Evangelos J. Giamarellos-Bourboulis,8 Willem H. Ouwehand,5,9 Jos W. M. van der Meer,2,7 Leo A. B. Joosten,4 Circa Wijmenga,7 Joost H. A. Martens,1 Ramnik J. Xavier,5,9,10 Colin Logie,1 Mihai G. Netea,2* Hendrik G. Stunnenberg1,7

IMMUNOGENETICS
There are 32 red cell blood group systems of man

The genetic basis of all blood groups has been resolved except for one rare group which is absent in 1 in 4000 individuals.

UN NOUVEAU FACTEUR SANGUIN « VEL »

Par Léon N. SUSSMAN, M.D. et Edward B. MILLER, M.D.

Au cours de l'étude sérologique d'une réaction transfusionnelle sévère, une agglutinine non encore décrite a été rencontrée. Dix mille spécimens de sang de groupe O, pris au hasard, furent testés, dont 4 seulement ne furent pas agglutinés. Cette fréquence inhabituelle (99,96 %) caractérise ce facteur comme l'agglutinogène le plus souvent présent dans la population blanche, surpassant les facteurs Cellano (99,8 %) [1], et hr” (e) (97 %) [2].

Le nom de ce facteur non décrit jusqu'ici a été choisi de façon à être en accord avec la nomenclature existante. Le nom du malade a favorisé ce dessein, et le facteur fut désigné comme « facteur Vel », son anticorps comme anti-Vel.
The genetic basis of all blood groups has been resolved except for one rare group which is absent in 1 in 4000 individuals.

Seventy-five genetic loci influencing the human red blood cell


rs1175550
AF = 77%

SMIM1
Thrombocytopenia with Absent Radii (TAR)

selective block of platelet formation and skeletal abnormalities

Bleeding because of low platelets

Missing radii and club hands

Shortened humerus and ulnae

Small shoulder girdle
The 5’UTR SNP and the intronic SNP lie in a regulatory element active in megakaryocytes and modify transcription factor binding
Inherited bleeding and platelet disorders (BPD)
864 rare cases enrolled across 12 clinical referral centres

Inclusion
- Abnormality of platelet
  - count / volume
  - morphology
  - function
- Bleeding of unknown aetiology
- Likelihood of being genetic

Exclusion
- Acquired causes
- Known inherited disorders

Phenotype

Genotype

Heterogeneity
Inherited bleeding and platelet disorders (BPD)

864 rare cases enrolled across 12 clinical referral centres

Phenotype

• Abnormality of platelet
  • count / volume
  • morphology
  • function
• Bleeding of unknown aetiology
• Likelihood of being genetic

Exclusion

• Acquired causes
• Known inherited disorders

The need to phenotype deeply

in order to reduce case heterogeneity and maintain power

Inclusion

Exclusion

Phenotype

Genotype

Heterogeneity
Human Phenotype Ontology annotation to cluster cases

Cases have been HPO coded by 12 clinical centres in the UK and overseas.
Human Phenotype Ontology annotation to cluster cases

Cases have been HPO coded by 12 clinical centres in the UK and overseas.
HPO coding enabled us to capture the phenotypic complexity of BPD cases. Neurological, immunological and skeletal disorders are highly over-represented.

Westbury *et al.* 2014 (submitted)
Computer based HPO-driven clustering of BPD cases across centres
it is hoped that this will maintain power of gene discovery

Phenotype Similarity

Westbury et al. 2014 (submitted)
WES vs WGS

HPS case identified by HPO clustering tested negative by WES but positive by WGS

Coverage – how often is a nucleotide seen

Hermansky Pudlak Syndrome 6 gene
Inherited bleeding and platelet disorders (BPD)
modern genomics added 11 genes since 2010
Inherited bleeding and platelet disorders (BPD)
modern genomics added 11 genes since 2010

Replicated

RBM8A
PLA2G4A
NBEAL2
TSPYL1
CYCS
GFI1B
ANKRD26
ACTN1
CHST14
COL1A1
NBEA

1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 X Y
Conclusions

1. Pilot projects have been essential to develop systems to achieve national scalability in 2015 for the 100,000 genomes

2. WGS provides excellent coverage of the virtual WES space and the regulome

3. Rare and common SNPs in regulatory elements are causative of rare conditions
Acknowledgements

UNIVERSITY OF CAMBRIDGE
Lucy Raymond
Willem H Ouwehand

Funding

Names in blue are at Cambridge University
Cambridge Team leaders in Bold and Italic