

# Disease gene discovery through genome sequencing and data exchange: perspectives for genetic diagnosis and drug development

20<sup>th</sup> International Symposium on  
Microsomes and Drug Oxidations  
Stuttgart, May 19<sup>th</sup>, 2014

H.H.Ropers (ropers@molgen.mpg.de)  
Max Planck Institute for Molecular Genetics, Berlin, Germany



# Genomics shifts focus to rare diseases

(from: E. Check-Hayden, Nature 2009)



Hugh Rienhoff sequenced family transcriptomes to try to diagnose his daughter Bea's genetic disease.

Since the introduction of NGS, no. of genes implicated in single gene disorders has doubled, from ~2000 in 2007 to >4000 in 2014

## ... because for most common diseases, GWAS has failed to identify genetic markers that are of diagnostic relevance

(from: H.H. Ropers, Am J Hum Genet 81:199, 2007)

- **Important role of non-genetic factors** (*life style [e.g., obesity], epigenetic and stochastic factors*)
- **Most risk alleles are evolutionarily young** *because of purifying selection (e.g., due to highly reduced reproductive fitness, as in ID, autism and SZ);* **for many common diseases, common risk factors cannot exist**
- **Complex disorders are often heterogenous, not multifactorial** (*many different gene defects cause the same phenotype; true for intellectual disability, but also blindness, deafness, autism, schizophrenia, epilepsy etc.*)

*While prominent supporters of the CD/CV hypothesis have changed their mind about GWAS...*

# The mystery of missing heritability: Genetic interactions create phantom heritability

Or Zuk<sup>a</sup>, Eliana Hechter<sup>a</sup>, Shamil R. Sunyaev<sup>a,b</sup>, and Eric S. Lander<sup>a,1</sup>

Broad Institute of MIT and Harvard, Cambridge, MA 02142; and <sup>b</sup>Genetics Division, Brigham and Women's Hospital, Harvard MA 02115

(Proc Natl Acad Sci USA, early ed., Jan. 2012)



'Phantom Heritability' Indicates Poor Predictive Value of Gene Tests

(Independent Science News, Jan. 10, 2012)

**The Emperor is naked!**

*... others go on undeterred.....*

# Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs

Cross-Disorder Group of the Psychiatric Genomics Consortium\*

VOLUME 45 | NUMBER 9 | SEPTEMBER 2013 NATURE GENETICS

*'Since the >370 authors of this article did not consider reduced reproductive fitness as a possible confounder of their meta-analysis, their conclusions about the genetic relationship between 5 psychiatric disorders should be interpreted with caution.'*

HHR, Letter to the Editor, submitted)

子曰：

Zi(3) yue(1):

眾惡之必察焉；眾好之，必察焉。

zhong(4)wu(4)zhi(1), bi(4)cha(2)yan(1); zhong(4)hao(3)zhi(1), bi(4) cha(2)yan(1)

**“What the masses condemn, examine critically; what the masses praise, examine critically.”**

(with thanks to Joe Terwilliger, 2011)

*And yet others intend to replace GWAS by genome sequencing to detect major genes, not markers, in common diseases- a good idea?*

## Whole-Exome Sequencing of 2,000 Danish Individuals and the Role of Rare Coding Variants in Type 2 Diabetes

Kirk E. Lohmueller,<sup>1,18,19</sup> Thomas Sparsø,<sup>2,18</sup> Qibin Li,<sup>3</sup> Ehm Andersson,<sup>2</sup> Thorfinn Korneliusson,<sup>4</sup> Anders Albrechtsen,<sup>5</sup> Karina Banasik,<sup>2</sup> Niels Grarup,<sup>2</sup> Ingileif Hallgrimsdottir,<sup>6</sup> Kristoffer Kiil,<sup>2</sup> et al

Am J Hum Genet 93:1072, 2013

‘.....Thus, we could reject a model for the genetic architecture of type 2 diabetes where rare non-synonymous variants clustered in a modest number of genes (fewer than 20) are responsible for the majority of disease risk.’

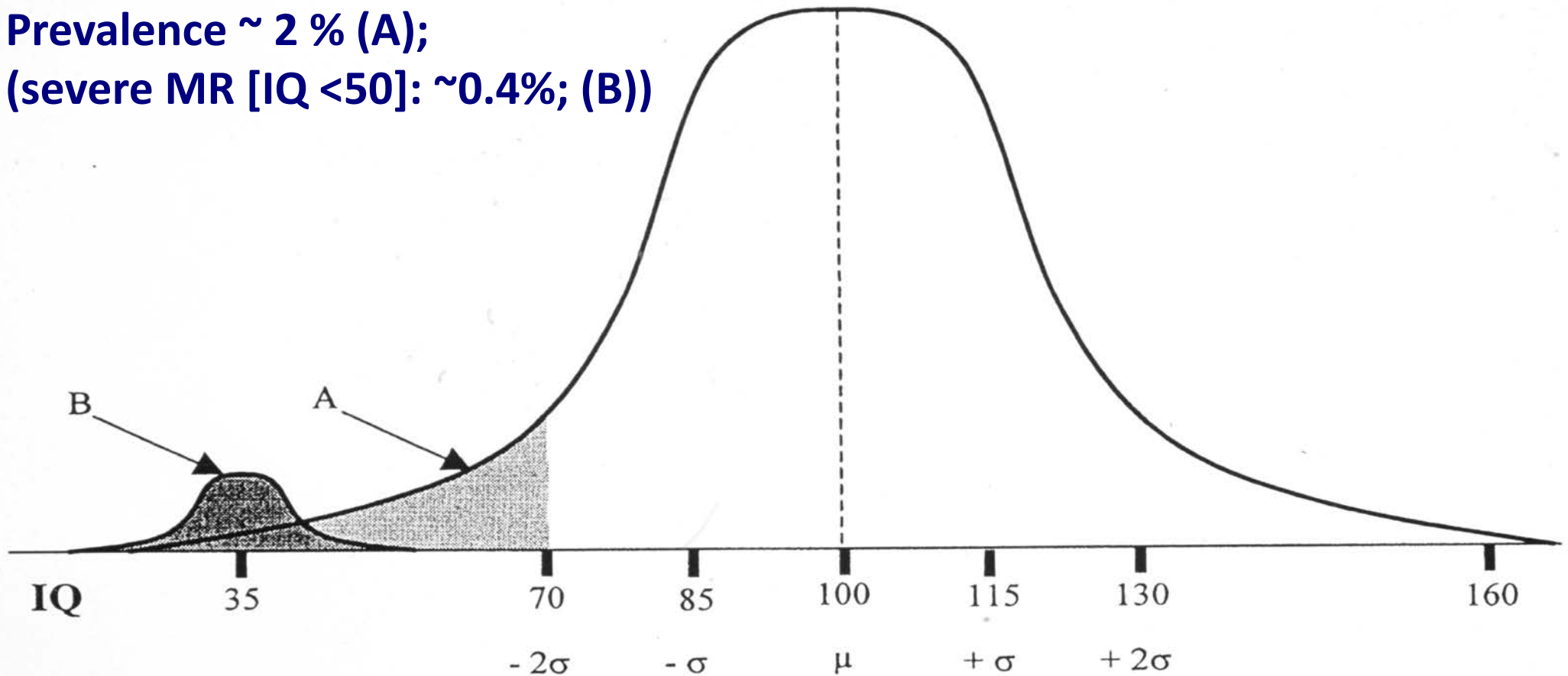
**Editorial:** ‘.....Although these findings might be seen as discouraging, they should be informative for the design and interpretation of future sequencing-based studies.’

(Ropers, MPIMG 2014)

# ID (IQ <70): a major socio-economic burden

*(lifetime costs 1-2M US\$; accounts for 11% of all disease-related costs in males, Roeleveld 1998, Gustavsson et al, 2011; most frequent reason for referral to genetic services, see also Yang et al, NEJM 2013)*

**Prevalence ~ 2 % (A);**  
**(severe MR [IQ <50]: ~0.4%; (B))**



# Severe intellectual disability (ID) in Western societies: mostly due to single genetic causes

20-25% chromosomal rearrangements (cytogenetically visible and submicroscopic ones)

10-12% X-linked forms (single gene defects or small del/dups, many already known) (Ropers & Hamel, Nat Rev Genet 2005)

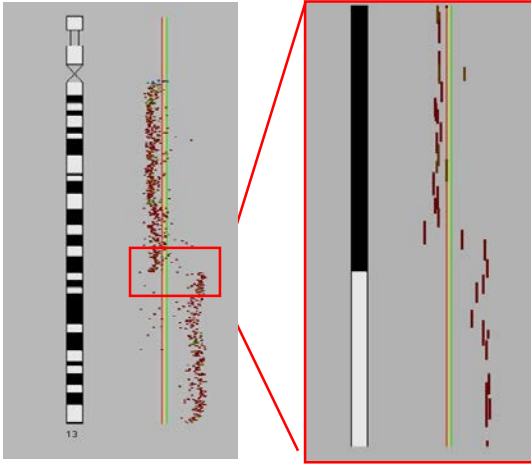
16-31% dominant *de novo* mutations (Rauch et al, Lancet 2012; de Ligt et al, NEJM 2012)

12-24% recessive defects (Musante & Ropers, TIGS 2013)

~ 25 (8-42) % other (non-exonic, di- or oligogenic, multifactorial, maternal alcohol abuse). Other non-genetic causes (e.g. perinatal hypoxia, infections, malnutrition) are rare (e.g., see Ropers, Annu Rev Genomics Hum Genet 2010)



# Array-based Comparative Genomic Hybridization (Array-CGH)



Potent method to detect copy number variants (CNVs, i.e., deletions or duplications) in clinically ,suspicious‘ patients

Covered by Statutory German Health Insurance, GOP11500, worth 1221.60 €, since Oct. 1st, 2013

Has revealed >200 recurrent, disease-linked CNVs, often with specific phenotypes. Account for  $\sim 1/3$  of the cases (mostly due to non-homologous pairing and recombination between flanking repeated sequences), but most disease-linked CNVs are non-recurrent and unique

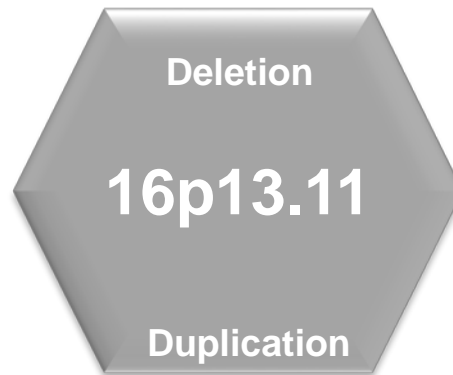
Some CNVs are **disease-causing**, others **predispose to disease**, with **high odds ratios (OR)**

# 16p13.11 deletions and duplications: among the most common CNVs seen in MR and related conditions (OR for ID: ~10)

**Intellectual Disability**  
Ullmann et al. 2007  
Hannes et al. 2009  
Mefford et al., 2009

**Schizophrenia**  
Ingason et al., 2011  
Magri et al., 2010  
Ikeda et al., 2010  
Kirov et al., 2009

**DISC1**



**NDE1**

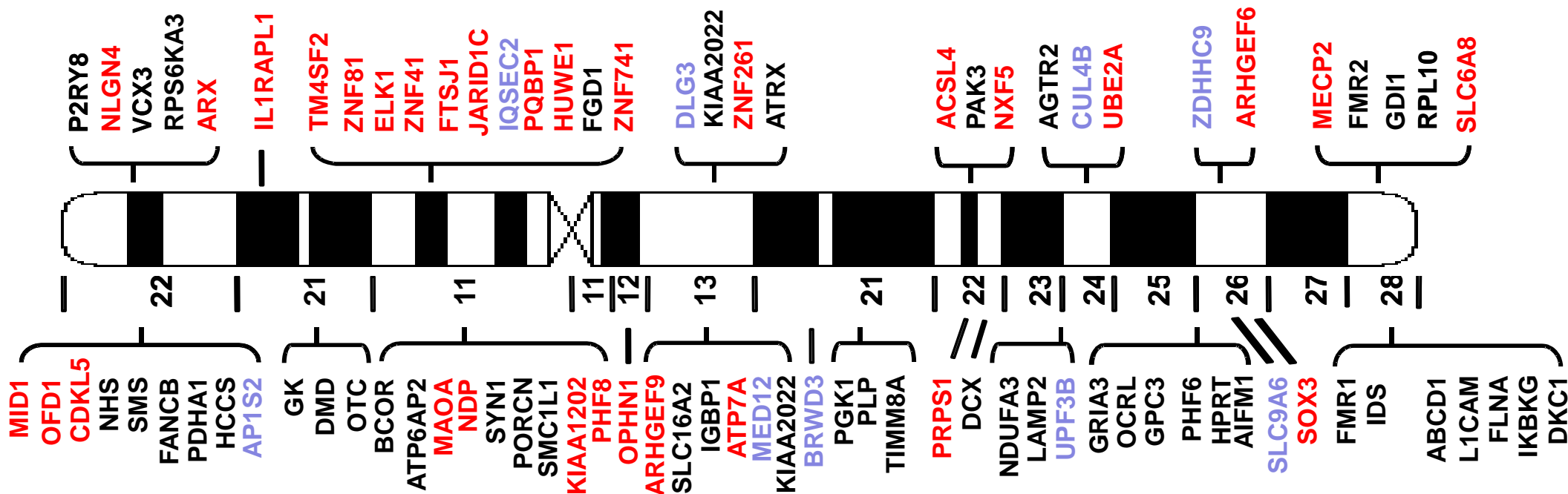
**ADHD**  
Williams et al., 2010

**LIS1 (PAFAH1B)**

**Epilepsy**  
Mefford et al., 2010  
Heinzen et al., 2010  
De Kovel et al., 2010

**Autism**  
Ullmann et al. 2007  
Hannes et al. 2009

# XLID genes identified until 2009

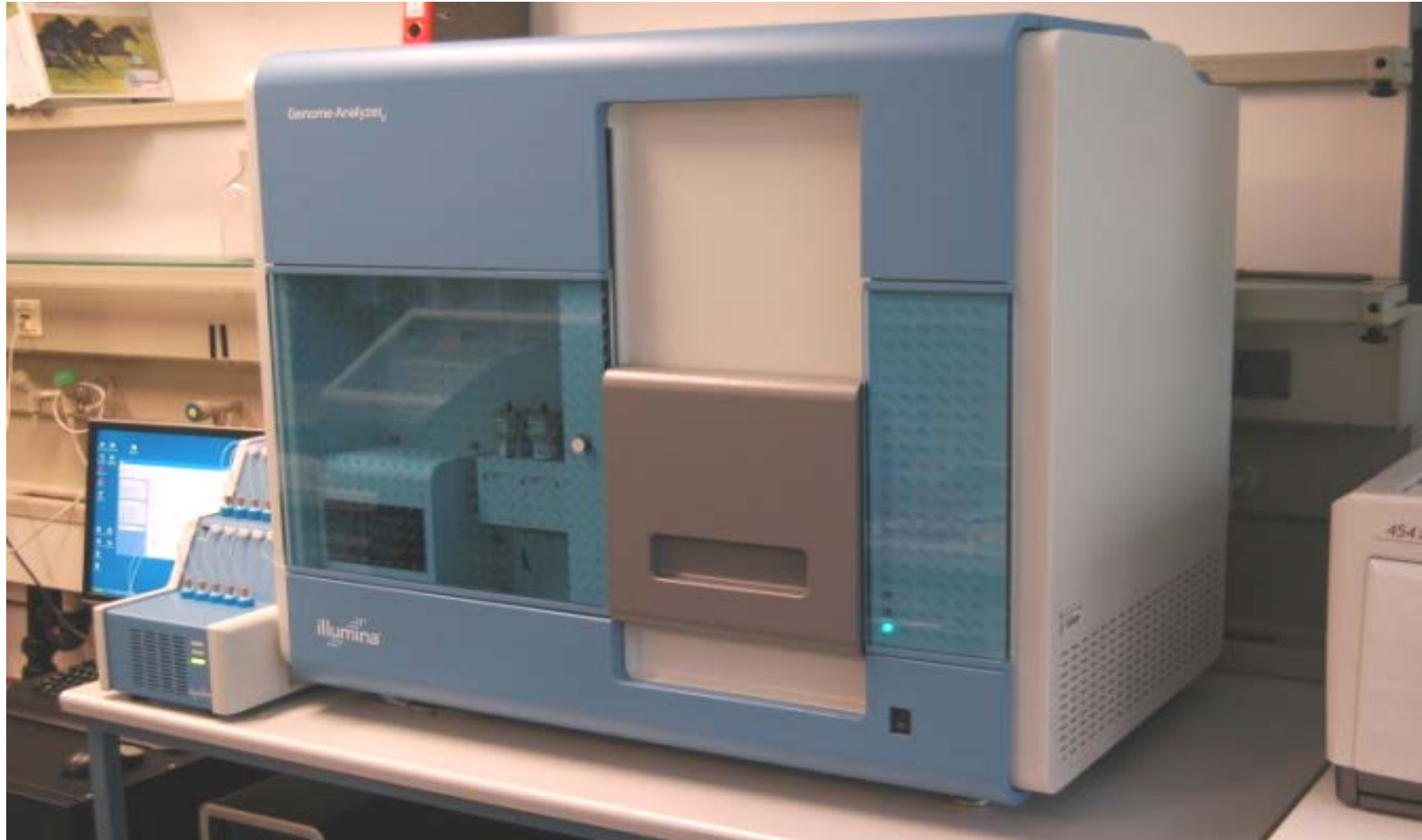


XLID genes identified by, or in collaboration with, the EURO-MRX Consortium (\*1995; de Brouwer et al., 2007)

IGOLD Consortium: Sanger sequencing of ~all X-chromosomal genes reveals 9 novel XLID genes (Tarpey et al, 2009)



# High-throughput - low cost sequencing revolutionizes genome research and genetic health care



**Solexa/Illumina Genome Analyzer**

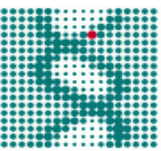
MPIMG Berlin, 2007(?)





**Exon sequencing in 248 unrelated families reveals 13 novel genes for X-linked intellectual disability; total no. of known XLID genes: 110**

- Up to 70% of XLID families carry mutations in these genes
- Missing mutations may hide in UTRs, introns and extragenic regulatory regions
- Total no. of ID genes: probably <155 ( based on no. of XLID genes identified by us and by Tarpey et al, Nature Genet. 2009)



## REPORT

---

### A Noncoding, Regulatory Mutation Implicates *HCFC1* in Nonsyndromic Intellectual Disability

Lingli Huang,<sup>1,2,12</sup> Lachlan A. Jolly,<sup>1,12</sup> Saffron Willis-Owen,<sup>1,11</sup> Alison Gardner,<sup>1,3</sup> Raman Kumar,<sup>3</sup> Evelyn Douglas,<sup>1</sup> Cheryl Shoubridge,<sup>1</sup> Dagmar Wieczorek,<sup>4</sup> Andreas Tzschach,<sup>5</sup> Monika Cohen,<sup>6</sup> Anna Hackett,<sup>7</sup> Michael Field,<sup>7</sup> Guy Froyen,<sup>8</sup> Hao Hu,<sup>5</sup> Stefan A. Haas,<sup>5</sup> Hans-Hilger Ropers,<sup>5</sup> Vera M. Kalscheuer,<sup>5</sup> Mark A. Corbett,<sup>1</sup> and Jozef Gecz<sup>1,3,9,10,\*</sup>

---

The American Journal of Human Genetics 91, 694–702, October 5, 2012

## ARTICLE

---

### An X-Linked Cobalamin Disorder Caused by Mutations in Transcriptional Coregulator *HCFC1*

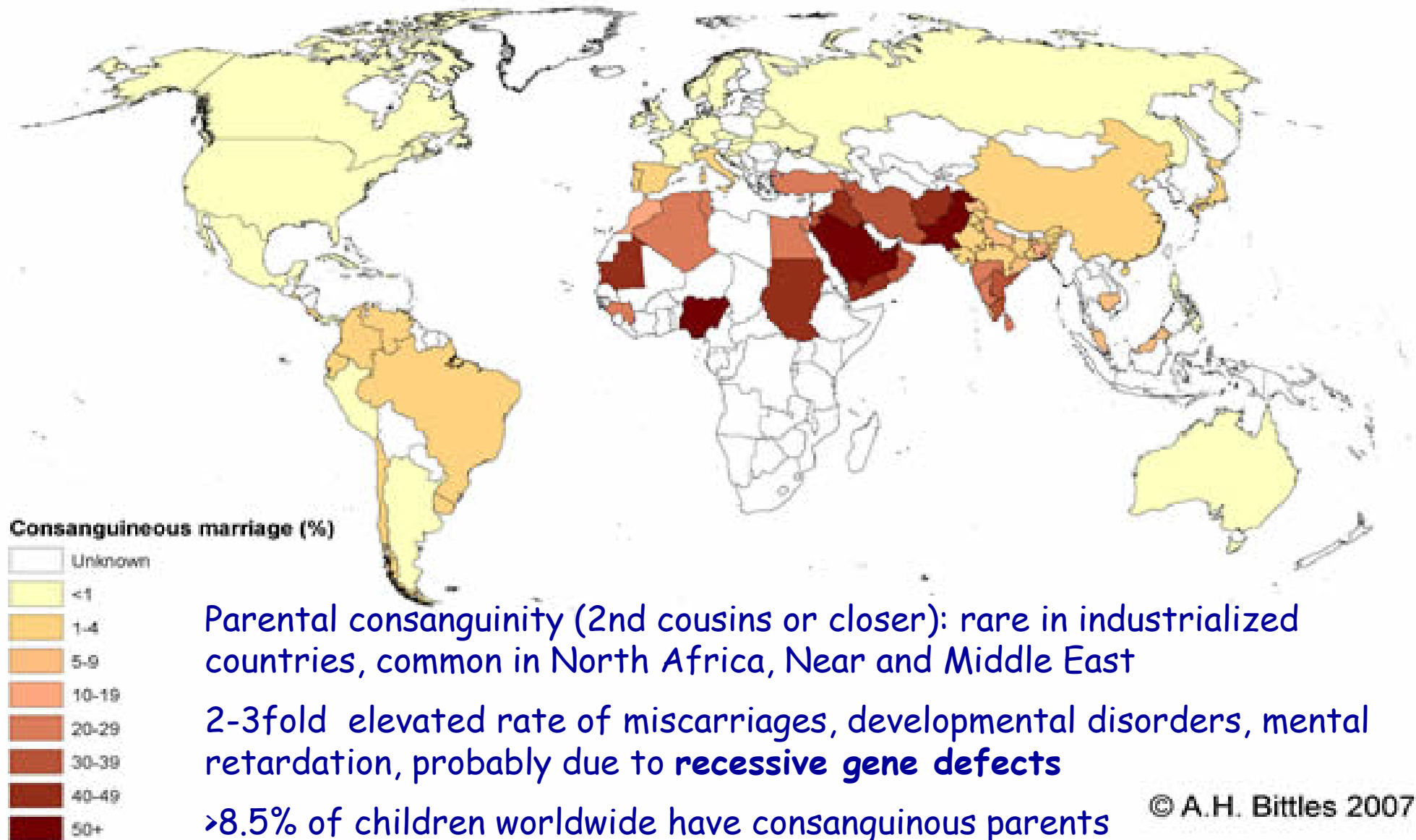
Hung-Chun Yu,<sup>1</sup> Jennifer L. Sloan,<sup>2</sup> Gunter Scharer,<sup>1,3,4,9</sup> Alison Brebner,<sup>5</sup> Anita M. Quintana,<sup>1</sup> Nathan P. Achilly,<sup>2</sup> Irimi Manoli,<sup>2</sup> Curtis R. Coughlin II,<sup>1,3</sup> Elizabeth A. Geiger,<sup>1</sup> Una Schneck,<sup>1</sup> David Watkins,<sup>5</sup> Terttu Suormala,<sup>6,7</sup> Johan L.K. Van Hove,<sup>1,3</sup> Brian Fowler,<sup>6,7</sup> Matthias R. Baumgartner,<sup>6,7,8</sup> David S. Rosenblatt,<sup>5</sup> Charles P. Venditti,<sup>2</sup> and Tamim H. Shaikh<sup>1,3,4,\*</sup>

(regulates gene for combined methylmalonic aciduria and homocystinuria on chr. 1p34)

---

The American Journal of Human Genetics 93, 506–514, September 5, 2013

# ID and other early-onset disorders are significantly more common in countries of the 'consanguinity belt'



# >450 consanguineous Iranian families (>200 with 3 and more patients)

This section displays several pedigree charts illustrating consanguineous relationships. Key features include:
 

- Generations I, II, III, and IV.
- Consanguineous unions (I-1 x I-2, II-1 x II-2, etc.).
- Shaded symbols representing affected individuals.
- Photographs of affected individuals, including a young boy (M192) and a young girl (M192).

~250 with non-syndromic MR:  
(no other clinical feature, FraX negative, normal karyotype)

This section displays several pedigree charts illustrating consanguineous relationships. Key features include:
 

- Generations I, II, III, and IV.
- Consanguineous unions (I-1 x I-2, II-1 x II-2, etc.).
- Shaded symbols representing affected individuals.
- Photographs of affected individuals, including a young girl (M159), a young man (M192), and a young man (M163).



78

# **Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci**

**Hossein Najmabadi · Mohammad Mahdi Motazacker · Masoud Garshasbi · Kimia Kahrizi · Andreas Tzschach · Wei Chen · Farkhondeh Behjati · Valeh Hadavi · Sahar Esmaeeli Nieh · Seyedeh Sedigheh Abedini · Reza Vazifehmand · Saghar Ghasemi Firouzabadi · Payman Jamali · Masoumeh Falah · Seyed Morteza Seifati · Annette Grüters · Steffen Lenzner · Lars R. Jensen · Franz Rüschen-dorf · Andreas W. Kuss · H. Hilger Ropers**

Received: 15 September 2006 / Accepted: 25 October 2006 / Published online: 21 November 2006  
© Springer-Verlag 2006

## Novel molecular defects underlying syndromic and non-syndromic ID

<b>Gene</b>	<b>Location</b>	<b>Function</b>	<b>Ethnicity</b>	<b>Reference</b>
<i>GRIK2</i>	<b>6q16.3</b>	Involved in the transmission of light signals from the retina to the hypothalamus, Involved in the maturation of microcircuits and network formation in brain areas	Iranian	Motazacker MM et al. Am J Hum Genet 2007; 81: 792–798
<i>TUSC3</i>	<b>8p22</b>	Putative Mg <sup>2+</sup> transporter, required for cellular Mg <sup>2+</sup> uptake. Indispensable for normal vertebrate embryonic development.	Iranian, French	Garshasbi M et al. Am J Hum Genet 2008; 82, 1158–1164
<i>VLDLR</i>	<b>9p24</b>	Part of the reelin signaling pathway, which is involved in neuroblast migration in the cerebral cortex and cerebellum	Iranian, Canadian, Turkish	Abbasi Moheb L et al. Euro J Hum Genet 2008; 16: 270–273
<i>TRAPPC9</i>	<b>8q24.3</b>	Enhancer of the cytokine-induced NF-(kappa)B signaling pathway, having an essential function in post mitotic neurons as opposed to neural progenitors	Iranian, Pakistani, Tunisian, Israeli	Mir A at al. Am J Hum Genet 2009; 85: 909-915
<i>SRD5A3</i>	<b>4q12</b>	Polyprenol reductase with a crucial role in N-linked protein glycosylation that is required for converting polyprenol to dolichol.	Iranian, Emirati, Turkish, Polish	Kahrizi K at al. Euro J Hum Genet 2011;19:115–117
<i>ZC3H14</i>	<b>14q31.3</b>	May contribute to control of gene expression in human cells through binding poly(A) RNA	Iranian	Pak CH et al., PNAS 2011
<i>ST3GAL3</i>	<b>1p34.1</b>	Transfers sialic acid to terminal positions on the carbohydrate groups of glycoproteins and glycolipids that are key determinants for a variety of cellular recognition processes	Iranian	Kuss AW et al., Am J Hum Genet 2011
<i>NSUN2</i>	<b>5p15.31</b>	RNA methyltransferase that methylates tRNAs, and possibly RNA polymerase III transcripts. May act downstream of Myc to regulate epidermal cell growth and proliferation	Iranian	Abbasi Moheb L et al., Am J Hum Genet 2012
<i>ZNF526</i>	<b>19q13.2</b>	Involved in transcriptional regulation, role in regulation of translation	Iranian	Abbasi Moheb L et al., ESHG meeting 2011

# Recent study quadruplicates the number of known genes for non-syndromic recessive mental retardation

## ARTICLE

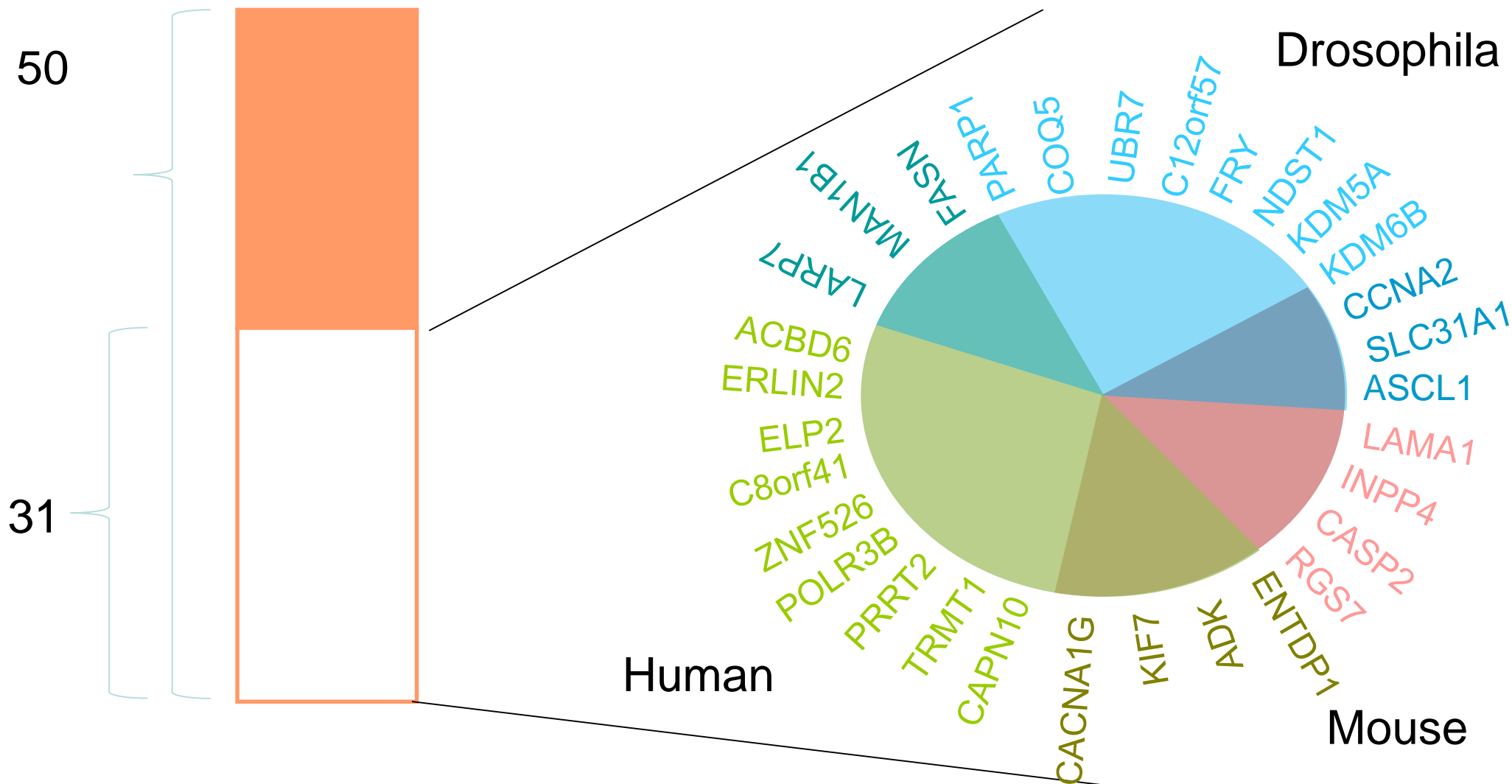
doi:10.1038/nature10423

# Deep sequencing reveals 50 novel genes for recessive cognitive disorders

Hossein Najmabadi<sup>1,2</sup>, Hao Hu<sup>3\*</sup>, Masoud Garshasbi<sup>1,3\*</sup>, Tomasz Zemojtel<sup>4</sup>, Seyedeh Sedigheh Abedini<sup>1</sup>, Wei Chen<sup>3,5</sup>, Masoumeh Hosseini<sup>1</sup>, Farkhondeh Behjati<sup>1</sup>, Stefan Haas<sup>4</sup>, Payman Jamali<sup>6</sup>, Agnes Zecha<sup>3</sup>, Marzieh Mohseni<sup>1</sup>, Lucia Püttmann<sup>3</sup>, Leyla Nouri Vahid<sup>1</sup>, Corinna Jensen<sup>3</sup>, Lia Abbasi Moheb<sup>1,3</sup>, Melanie Bienek<sup>3</sup>, Farzaneh Larti<sup>1</sup>, Ines Mueller<sup>3</sup>, Robert Weissmann<sup>3</sup>, Hossein Darvish<sup>1</sup>, Klaus Wrogemann<sup>3,7</sup>, Valeh Hadavi<sup>2</sup>, Bettina Lipkowitz<sup>3</sup>, Sahar Esmaeeli-Nieh<sup>3</sup>, Dagmar Wiczorek<sup>8</sup>, Roxana Kariminejad<sup>2</sup>, Saghar Ghasemi Firouzabadi<sup>1</sup>, Monika Cohen<sup>9</sup>, Zohreh Fattahi<sup>1</sup>, Imma Rost<sup>10</sup>, Faezeh Mojahedi<sup>11</sup>, Christoph Hertzberg<sup>12</sup>, Atefeh Dehghan<sup>13</sup>, Anna Rajab<sup>14</sup>, Mohammad Javad Soltani Banavandi<sup>1</sup>, Julia Hoffer<sup>3</sup>, Masoumeh Falah<sup>1</sup>, Luciana Musante<sup>3</sup>, Vera Kalscheuer<sup>3</sup>, Reinhard Ullmann<sup>3</sup>, Andreas Walter Kuss<sup>3,†</sup>, Andreas Tzschach<sup>3</sup>, Kimia Kahrizi<sup>1</sup> & H. Hilger Ropers<sup>3</sup>

Nature 478:57-63, Oct. 6th, 2011

**Follow-up:** validation of ARID candidate genes; WGS in families without plausible variants; WES in second cohort (>300 consanguineous ARID families: no sign of saturation!)



(courtesy L. Musante and 'Cougar' Hao Hu, MPIMG Berlin, 02/2014)

# Next Generation Sequencing: three different flavors

## -Panel Sequencing (PS)

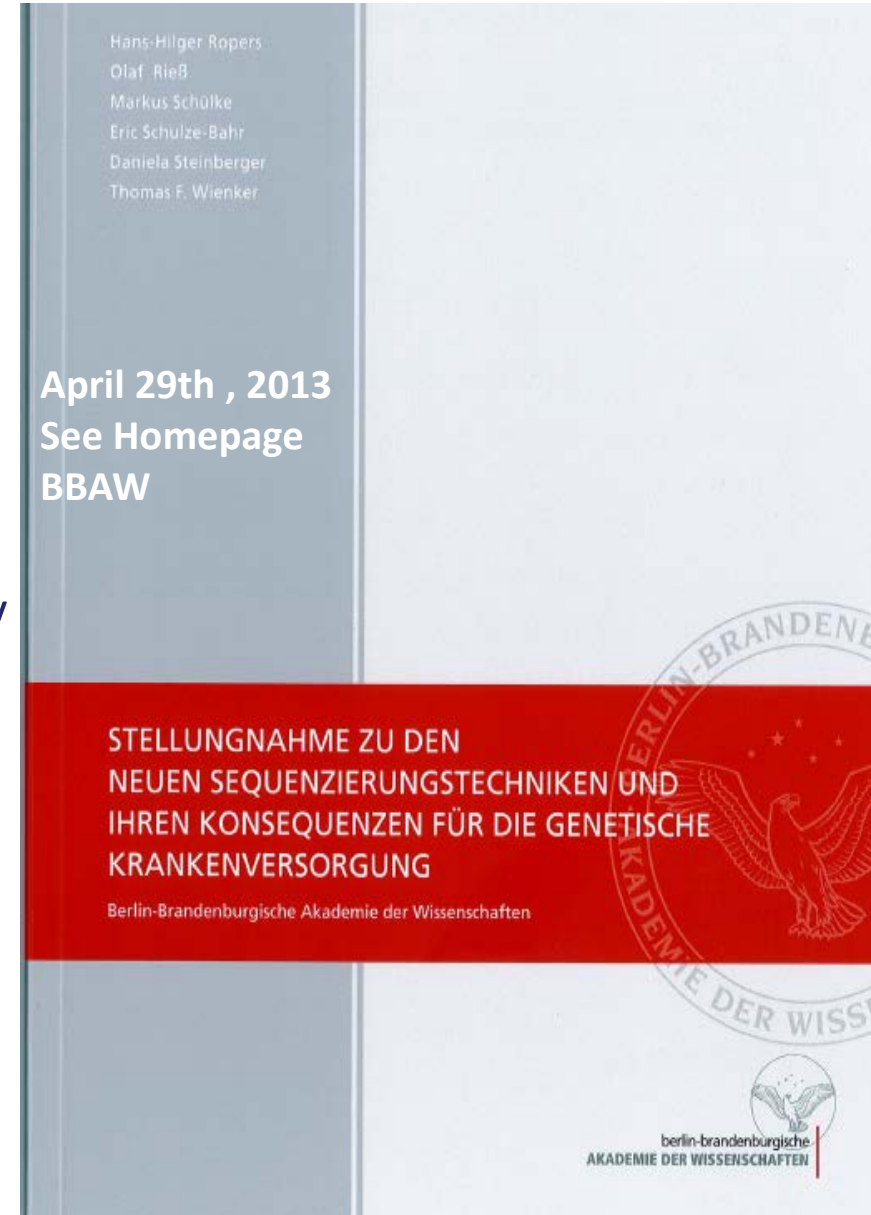
Several dozen to several thousand genes, custom-made enrichment kits, very high coverage, very high sensitivity and specificity, low costs (mostly < 1000 €), **no ,unsolicited findings‘**

## -Whole Exome Sequencing (WES)

Exonic sequences of all ~22k protein-coding genes, Commercial kits available, coverage/sensitivity/specificity lower than for PS, higher costs (>1500 €); not confined to known genes, Trio-WES → novel mutations.

## -Whole Genome Sequencing (WGS)

Includes non-coding parts of all genes and intergenic regions; no enrichment - more even coverage. Attractive commercial services, but time-consuming. Rapid sequencing protocols for NICU



# Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders

(Yang et al, NEJM 369:1502-1511, 2013)

,First 250 consecutive probands for whom referring clinicians ordered whole-exome sequencing..... **Approximately 80% were children with neurologic phenotypes.**

62 of these carried **mutations in known disease genes** 'that were highly likely to be causative' (diagnostic success rate 25%), including 4 patients with two non-overlapping disorders.

66 diseases in total, **36 autosomal dominant (54,5%), 20 autosomal recessive (30.3%), 10 X-linked (15.2%).**

# Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing

Callum J. Bell,<sup>1\*</sup> Darrell L. Dinwiddie,<sup>1,2\*</sup> Neil A. Miller,<sup>1,2</sup> Shannon L. Hateley,<sup>1</sup>  
Elena E. Ganusova,<sup>1</sup> Joann Mudge,<sup>1</sup> Ray J. Langley,<sup>1</sup> Lu Zhang,<sup>3</sup> Clarence C. Lee,<sup>4</sup>  
Faye D. Schilkey,<sup>1</sup> Vrunda Sheth,<sup>4</sup> Jimmy E. Woodward,<sup>1</sup> Heather E. Peckham,<sup>4</sup>  
Gary P. Schroth,<sup>3</sup> Ryan W. Kim,<sup>1</sup> Stephen F. Kingsmore<sup>1,2†</sup>

*(Sci. Transl. Med. 2011)*

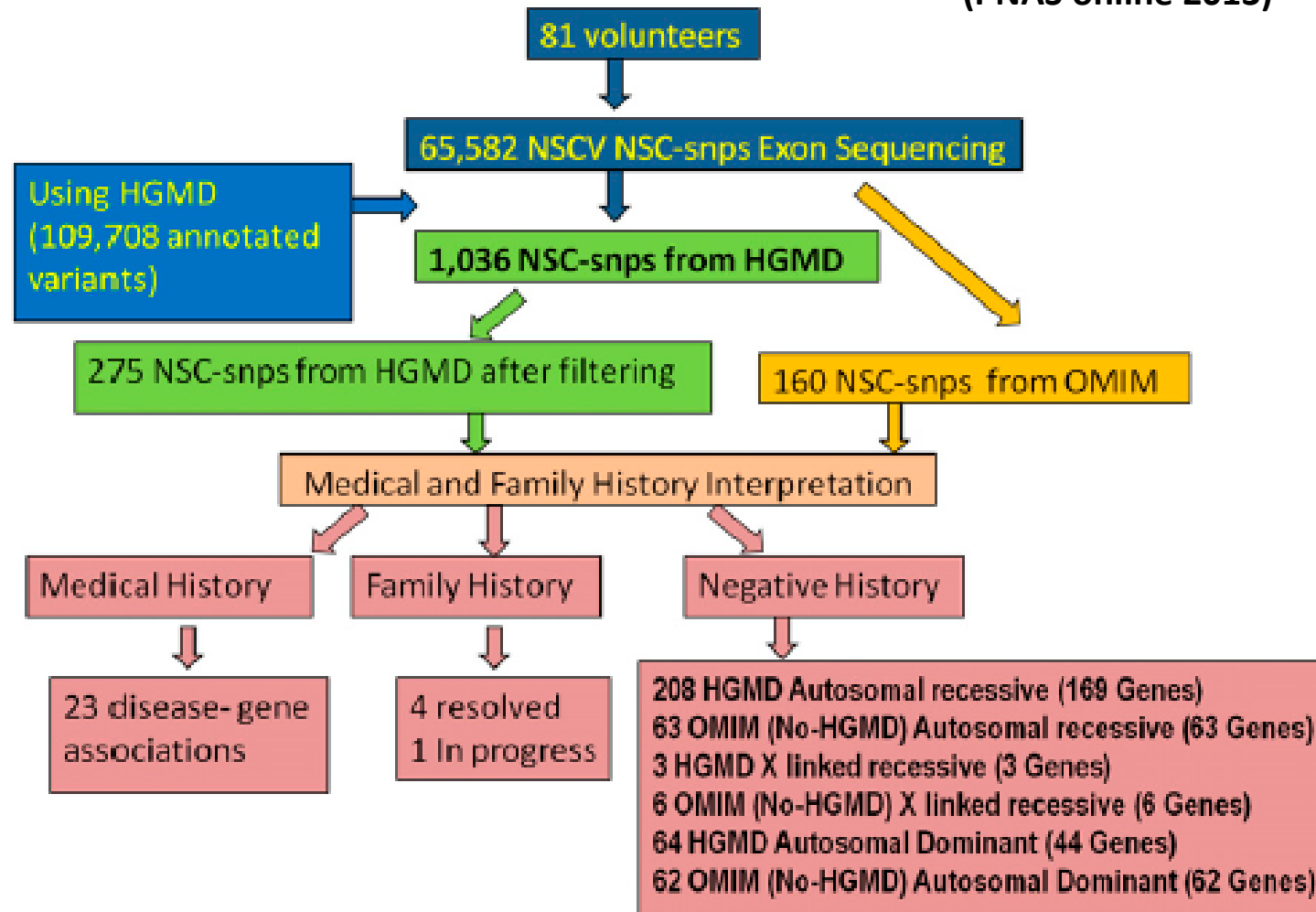
*,Healthy individuals harbour on average 2.8 known recessive severe childhood disease mutations‘*

*,27% of mutations cited in the literature are common polymorphisms or misannotated‘*

# Personalized genomic disease risk of volunteers

Manuel L. Gonzalez-Garay<sup>a,1</sup>, Amy L. McGuire<sup>b</sup>, Stacey Pereira<sup>b</sup>, and C. Thomas Caskey<sup>c,1</sup>

(PNAS online 2013)



**„Diagnostic‘ WES in 81 healthy volunteers reveals 271 presumably pathogenic autosomal recessive variants (3.3 per individual)**



# Promoting NGS as health care tool: partnering with parents and media



In isolated patient with syndromic ID, WGS reveals *de novo* stop mutation in 'novel' ASXL3 gene; no other plausible change in entire genome

2nd patient identified by Baylor colleagues



Subject 2



Subject 3



Subject 4



Subject 5



Subject 6



Subject 7



Subject 5 died  
23 h after birth



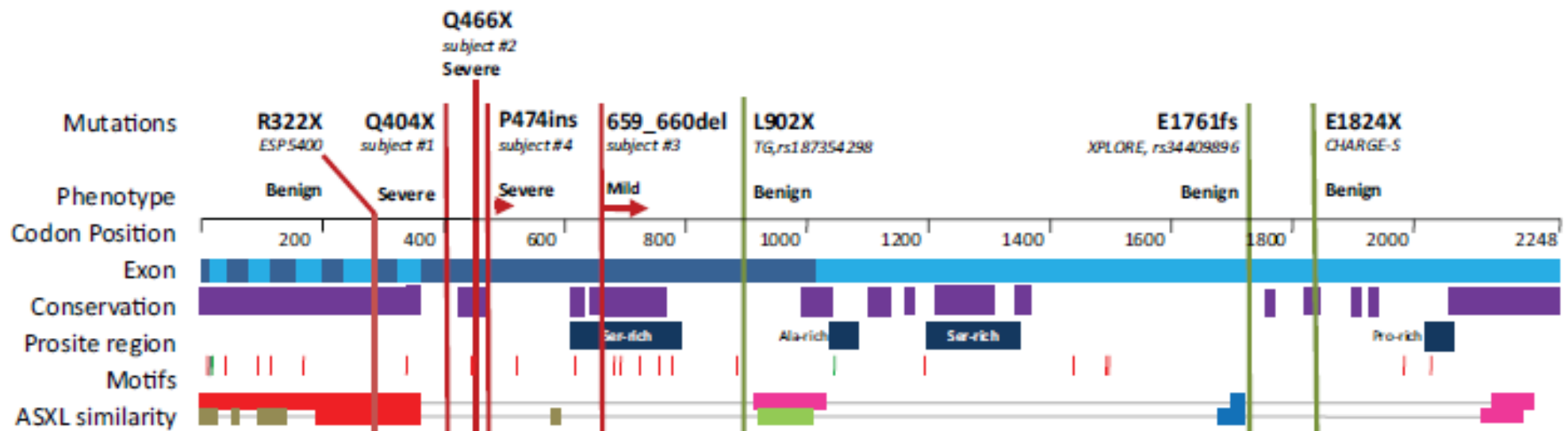
Patients with Bohring-Opitz syndrome have *de novo* truncating mutations in related ASXL1 gene - but not all... (Hoischen et al, Nature Genet 2011)

# *De novo* truncating mutations in *ASXL3* are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome

Matthew N Bainbridge<sup>1,2</sup>, Hao Hu<sup>3</sup>, Donna M Muzny<sup>1</sup>, Luciana Musante<sup>3</sup>, James R Lupski<sup>1,2,4,5</sup>, Brett H Graham<sup>2,5</sup>, Wei Chen<sup>3,6</sup>, Karen W Gripp<sup>7</sup>, Kim Jenny<sup>7</sup>, Thomas F Wienker<sup>3</sup>, Yaping Yang<sup>2</sup>, V Reid Sutton<sup>2,5</sup>, Richard A Gibbs<sup>1,2\*</sup> and H Hilger Ropers<sup>3\*†</sup>

(Genome Med 2013)

(OMIM: ,Bainbridge-Ropers syndrome')



**Figure 2 Known nonsense mutations in *ASXL3*.** Known nonsense mutations are shown (dotted line) across the *ASXL3* gene, and are given with the phenotypic effect and DNA source. Amino acid position is listed (top) with exon position (light/dark blue). Regions of high vertebrate conservation are shown (purple) along with Prosite regions (dark blue), and conserved and predicted motifs (nuclear localization in black, ASXL-specific motifs in green, phosphorylation sites in red) amino acid similarity between *ASXL3* and *ASXL1* (top) and *ASXL2* (bottom) (highest similarity to lowest: red, pink, green, blue, brown). Adapted from UCSC Genome browser, ENSEMBL, eukaryotic linear motif server, and NCBI-BLASTP.

## MEDICINE

# Data barriers limit genetic diagnosis

*Tools for data-sharing promise to improve chances of connecting mutations with symptoms of rare diseases.*

BY ERIKA CHECK HAYDEN

# Frankfurter Allgemeine

ZEITUNG FÜR DEUTSCHLAND

22.11.2013

Von H.-Hilger Ropers

## Wer hat Deutungshoheit über das Genom?

Zwischen Kommerz und Patientenfürsorge: Genetische Diagnostik sollte an Unikliniken angebunden werden und nicht Ärzten und Krankenkassen überlassen werden.

\* Myriad Genetics loses patents for breast cancer diagnosis, but market value of company soars – because of its publicity campaign highlighting the importance of its proprietary database for assessing the disease relevance of BRCA 1 and 2 mutations

\* BGI, world-largest genome sequencing provider, announces plans to sequence 1 million patients („we will sequence the world“). Motivation? Establishment of similar database, but covering all human genes? Commercial intentions?

**South China Morning Post**

December 16<sup>th</sup>, 2013

# Biotech firm dreams of changing drugs world

Beijing Genomics Institute hopes to use recent acquisition to build DNA database

RayChan [ray.utchan@scmp.com](mailto:ray.utchan@scmp.com)

A small biotech start-up company based in Shenzhen has had a dream of developing an intriguing technology that might change the landscape of the drug-making industry in China and the world. Wang Jun, a biologist-turned-entrepreneur is behind the Beijing Genomics Institute (BGI), a bioinformatic, genomics, and technology company. It made a significant step towards becoming a personalized drug-maker after completing the US\$119 million acquisition of Complete Genomics, a US-listed human genome sequencing company.

In an interview with the *South China Morning Post*, 37-year-old Wang said he wanted to build up "a big database of 100 million DNA combinations in human beings through leveraging Complete Genomics' existing intellectual property rights and devices". After establishing the genetic sequence database, he said, many diseases could be cured based on the findings. Wang, who received his doctoral degree from Peking University, said one of the challenges that weighed on the acquisition process was the loss making US company's continuing patent lawsuit hiatus, though cross-border legal experts have helped ease the legal uncertainties.

***Medical genome sequencing and comprehensive databases:  
Key to genetic diagnosis and drug development***

# **‘Illumina’s new HiSeq XTen machine could slash cost of sequencing your genome to \$1,000’ (January 14<sup>th</sup>, 2014)**



- **Novel sequencing system (price US\$ 10 million), 4,5x reduction of genome sequencing costs**
- **Enormous capacity (18.000 genomes/year): ideal for National Genome Sequencing Center(s)**
- **Game changer for diagnosis of rare diseases: ‚sequencing first‘-strategy takes days or weeks, error-prone ‚syndromologist-first‘-protocol takes ~7 years (in D)**
- **Storage of all ‚novel‘ data in central database, waiting for ‚second case‘**
- **Stakeholders: Healthcare, Life Science, Pharmaceutical Industry – and Governments**

# ~4000 (out of 15.000?) genes implicated in single gene disorders: miles to go before we sleep.....

- ,Genomics England'(2013-2018): governmental disease-focused 100k genome project (rare diseases, cancer, infectious diseases); a role model?
- BGI/Complete Genomics plans to perform diagnostic whole-genome sequencing in 1 million patients (Huanming Yang, ASHG Boston, 2013) ; Wang Jun's plans for largest database, aim: drug development (South China Morning Post, Dec 16th, 2013)
- **introduction of deep seqencing as first-line diagnostic test!** One central WGS facility with central database serving/run by all medical genetics centers at university hospitals, which in turn serve as reference centers for other geneticists and specialists

## Take-home messages

- Common diseases are very heterogeneous → studying large cohorts to identify major genetic determinants of disease does not make much sense. Most families → different genetic entities, to be studied separately
- Genetic defects that do not cause, but predispose to disease (e.g., 16p13.11 del/dups, truncating ASXL3 mutations in Bainbridge-Ropers syndrome and healthy individuals) are likely to be common; major problem for genetic counseling
- WGS for 1000\$: implementation as 'one size fits all', first-line diagnostic test for all patients with unclear, possibly genetic conditions will immediately solve ~30% of the cases and dramatically shorten time to diagnosis
- Strong arguments for concentration at central facility serving all medical centers (quality and cost control; one central database for variants of unknown significance; prerequisite for identification of novel disease-causing variants and disease genes)





# ‘Je moet niet bang zijn voor je eigen genoom’

*Hans-Hilger Ropers, NRC Handelsblad, October 2012*



# Improving genome understanding

*The cost and accuracy of genome sequencing have improved dramatically. George Church asks why so few people are opting to inspect their genome.*

*Nature, October 2013*



## Why not elucidate all monogenic disorders?

Because it is doable, affordable and immediately beneficial for the relevant families (family planning, prevention). Costs can be estimated fairly accurately, much lower than previous GWAS. Splitting costs between Ministries of Health and Research should be an option.

In the medical genome sequencing era, elucidation of genetic defects is possible without large affected families, which are rare both in China and in Germany.

A gene by gene approach like this has a defined start, a defined end, and most involved will live to see its results.

Still, complexity will pop up everywhere (e.g., incomplete penetrance of many 'disease-causing defects'; already observed in CNVs; why are males more often affected than females; different manifestations in different populations). Mutations in non-coding sequences (introns, UTRs, non-coding RNA genes, intergenic regions) requiring particularly 'telling' specific phenotypes for their identification. Along the same lines: gene-gene interactions, gene regulation.

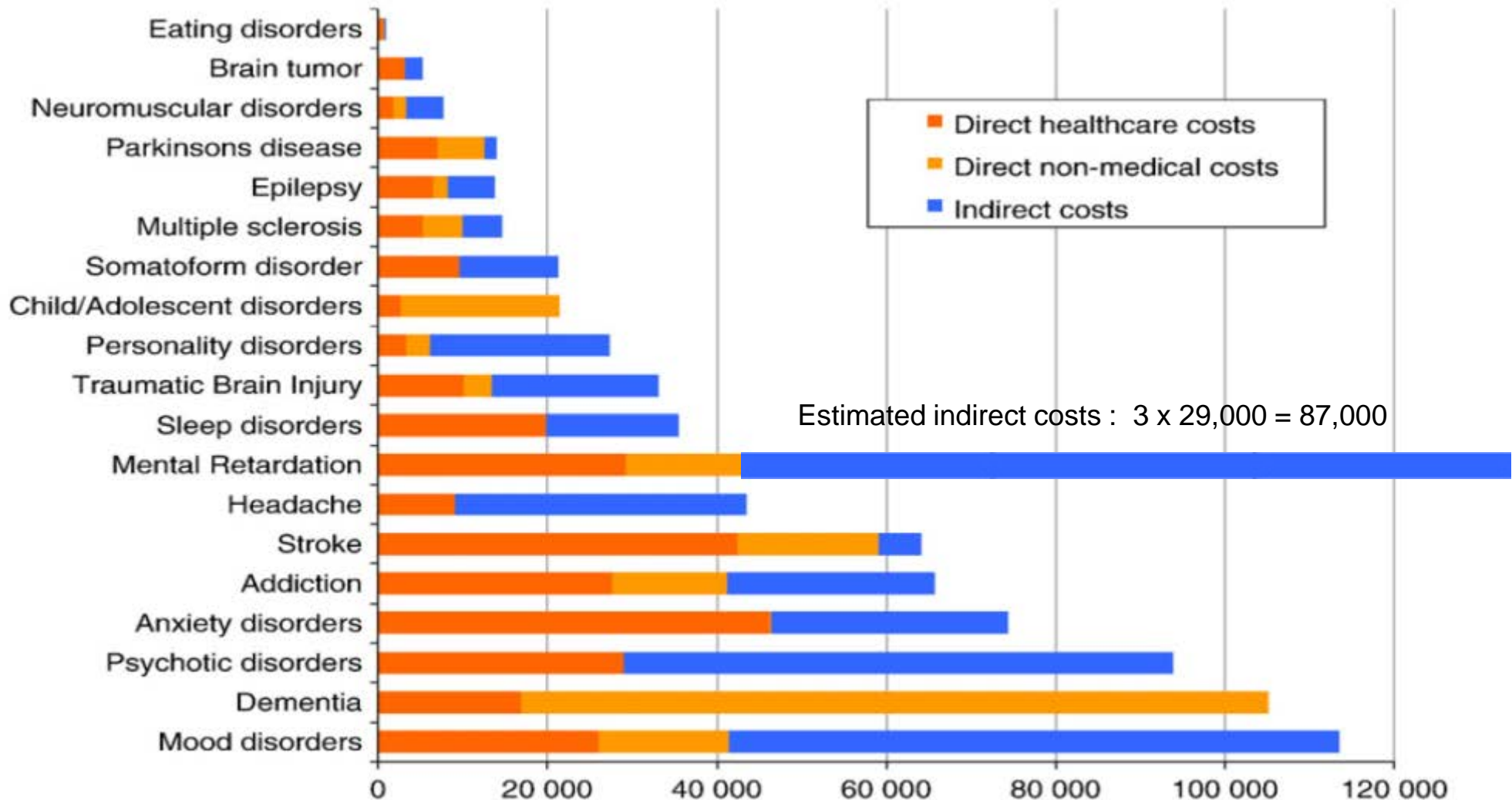
Phenotyping is much easier in humans than e.g., in mice (where genes are being knocked out in systematic fashion); requires no additional effort; model organisms have often different phenotypes or are not affected at all

It will instill blood into Personal Medicine, predictive diagnosis is presently confined to monogenic disorders, cancer and pharmacogenetic traits;

Will shed light on the function of many human genes, often with immediate consequences for drug development, unlike associated markers for complex disorders which have no functional significance.

Will also contribute to understanding of many common diseases with monogenic components. Often life style is more important than genetic predispositions, and in view of the empirical evidence, studying large cohorts by WGS is at least risky.

**Brain disorders cost Europe almost €800 billion (US\$1 trillion) a year — more than cancer, cardiovascular disease and diabetes put together.**



**Figure 3** Total cost by disorder and type of cost (€PPP million, 2010), all disorders.

after Gustavsson A et al. (2011) , Eur Neuropsychopharmacol 21(10): 718 –779, Epub 2011 Sep 15