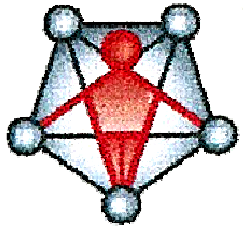
Pan-European Biobanking and Biomolecular Resources Research

Prof dr Gertjan van Ommen

Centre for Medical Systems Biology

Leiden University Medical Centre, NL

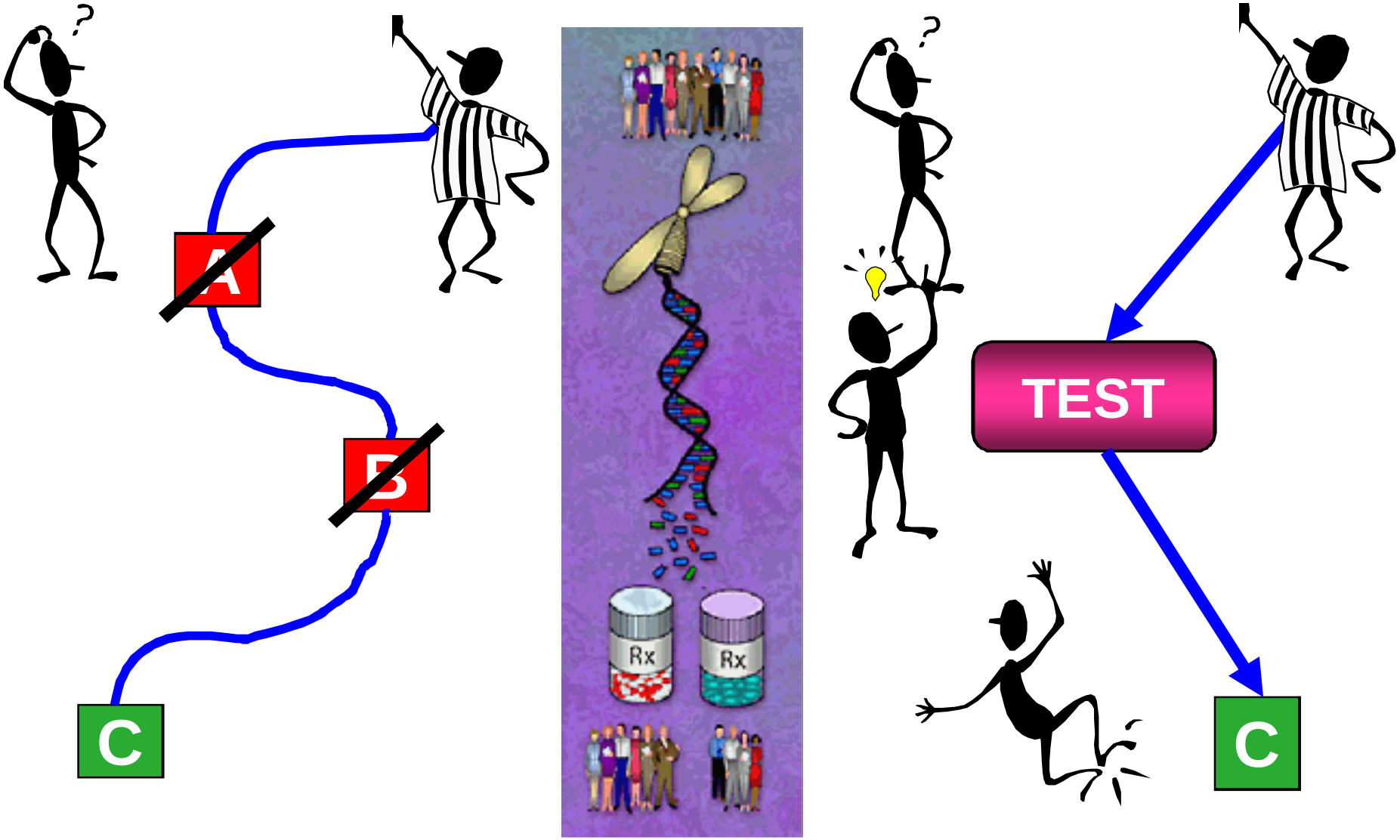


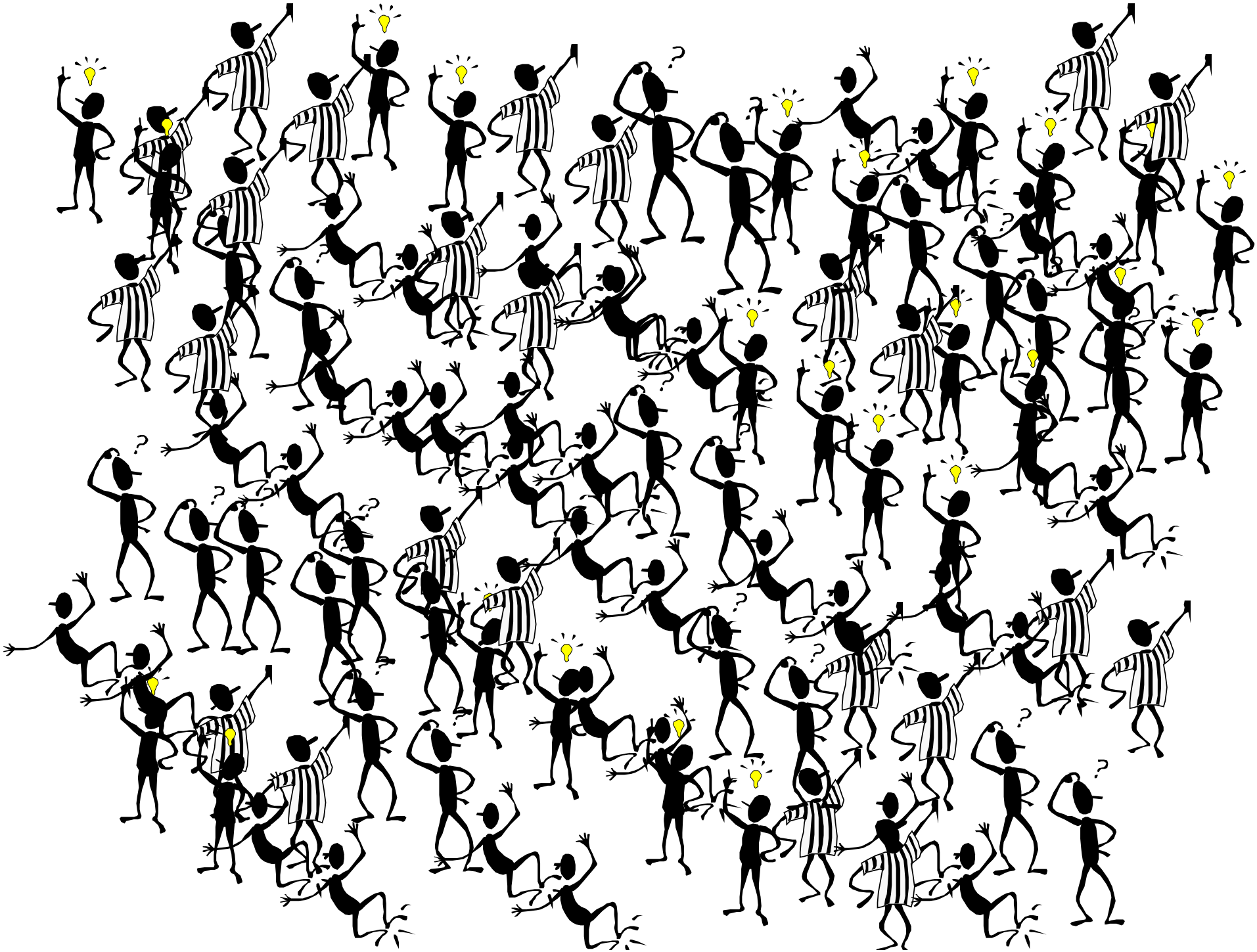
PHARMACOGENOMICS:

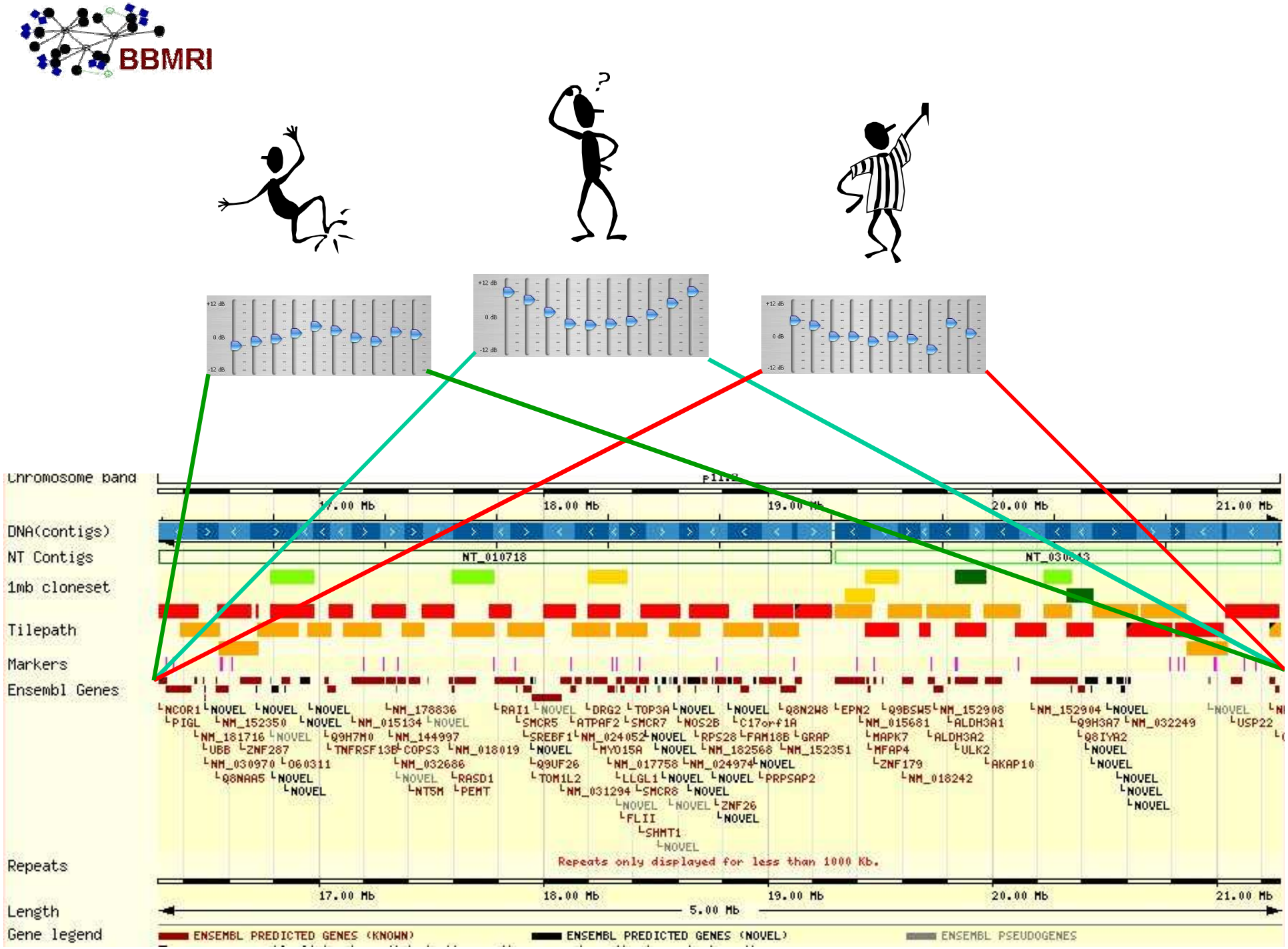
Genomics and biobank research into drug treatment outcomes

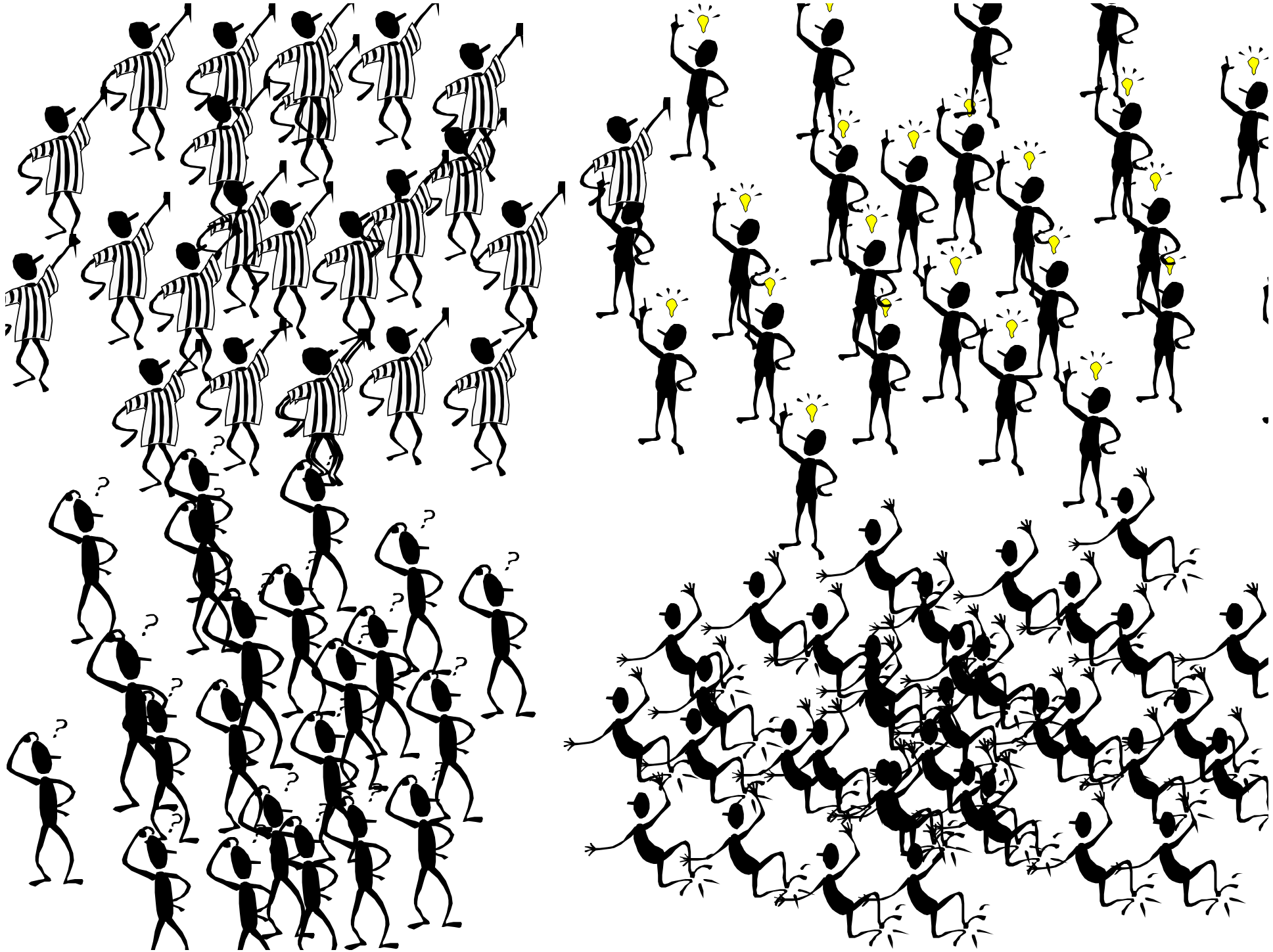
NOW: TRIAL AND ERROR

LATER: PERSONALIZED MEDICINE






BBMRI
Chromosome band
17.00 Mb
18.00 Mb
19.00 Mb
20.00 Mb
21.00 Mb
DNA(contigs)
NT Contigs
NT_010718
1mb cloneset
Tilepath
Markers
Ensembl Genes
Repeats
Repeats only displayed for less than 1000 Kb.
Length
5.00 Mb
Gene legend
ENSEMBL PREDICTED GENES (KNOWN)
ENSEMBL PREDICTED GENES (NOVEL)
ENSEMBL PSEUDOGENES



Considerations

Complex diseases like Alzheimer's, asthma, arthritis, cancer, cardiovascular diseases, diabetes, hypertension, obesity, Parkinson, and psychiatric disorders are the number one cause of disease burden (77%) and deaths (88%) across Europe.

Large number of small, often additive effects from genetic predisposition, lifestyle and the environment.

➔ Comparing large numbers of affected and unaffected individuals

Europe

Long tradition of excellence in education, research, medical care and registries

Understanding the etiology of complex diseases: international coordination and collaboration of biobanked data and samples essential requisite

➔ ***European biobanks a great asset and one of the few competitive advantages vs the US and Japanese research communities.***

COVERAGE

New: UK Biobank, Iceland, Estonia, Twin registries, NL Pearl String Project, many other emerging initiatives

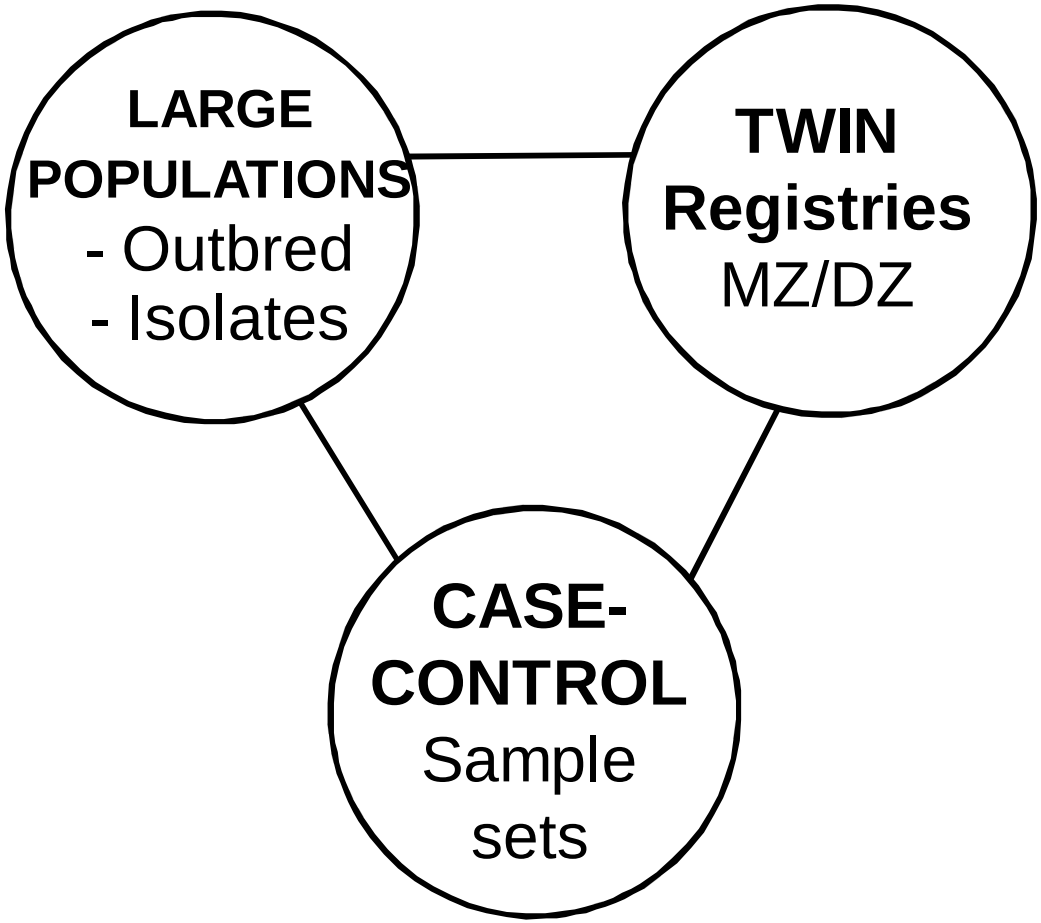
Existing: Systematic collections, for decades, through national health care systems, hundreds of millions of samples - *a major investment!*

Formats:

Population-based health surveys and biobanks
Population isolates
Twin registries (MZ: Environment – DZ: Genes)
Case-contol disease cohorts
Not only diseases, also healthy ageing:
positive reference for disease patterns

BIOBANKING CONNECTIVITY

Maximizing the impact of biomedical genomics



INTEGRATIVE GENOMICS

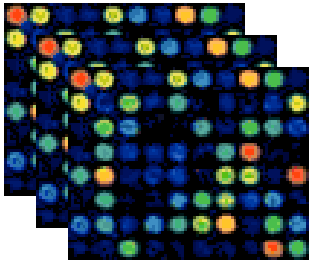
Systems Biology:

Toward multidimensional analytical tools

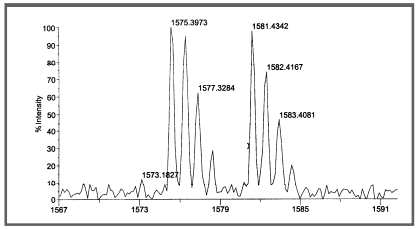
Human and animal model systems:

- tissues
- blood cells
- plasma
- urine
- cerebrospinal fluid

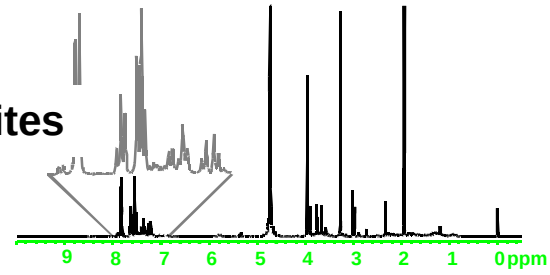
Genes

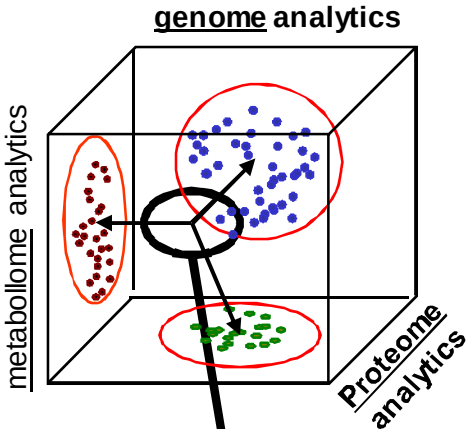


Proteins



Metabolites





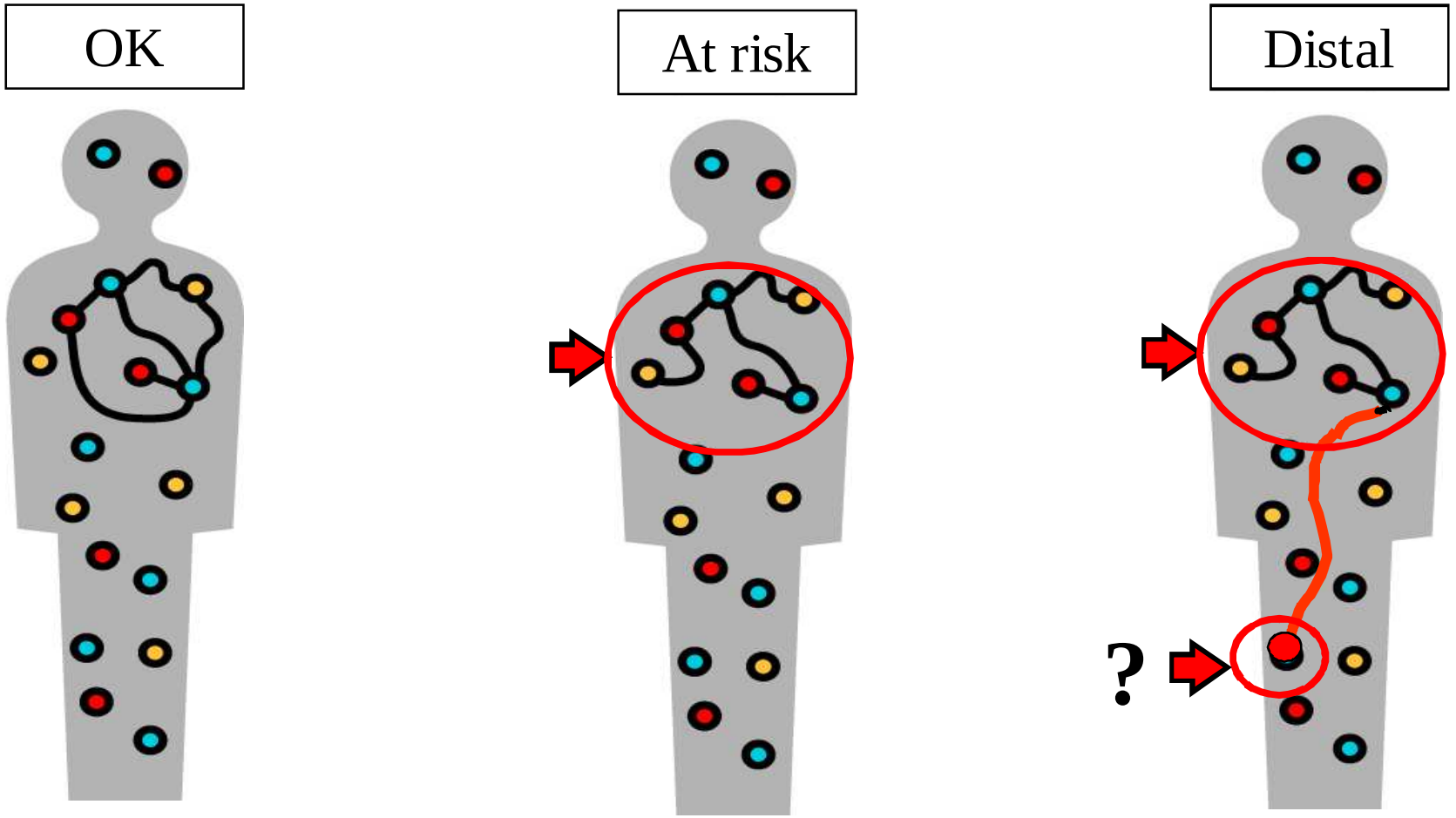
Integrative Genomics

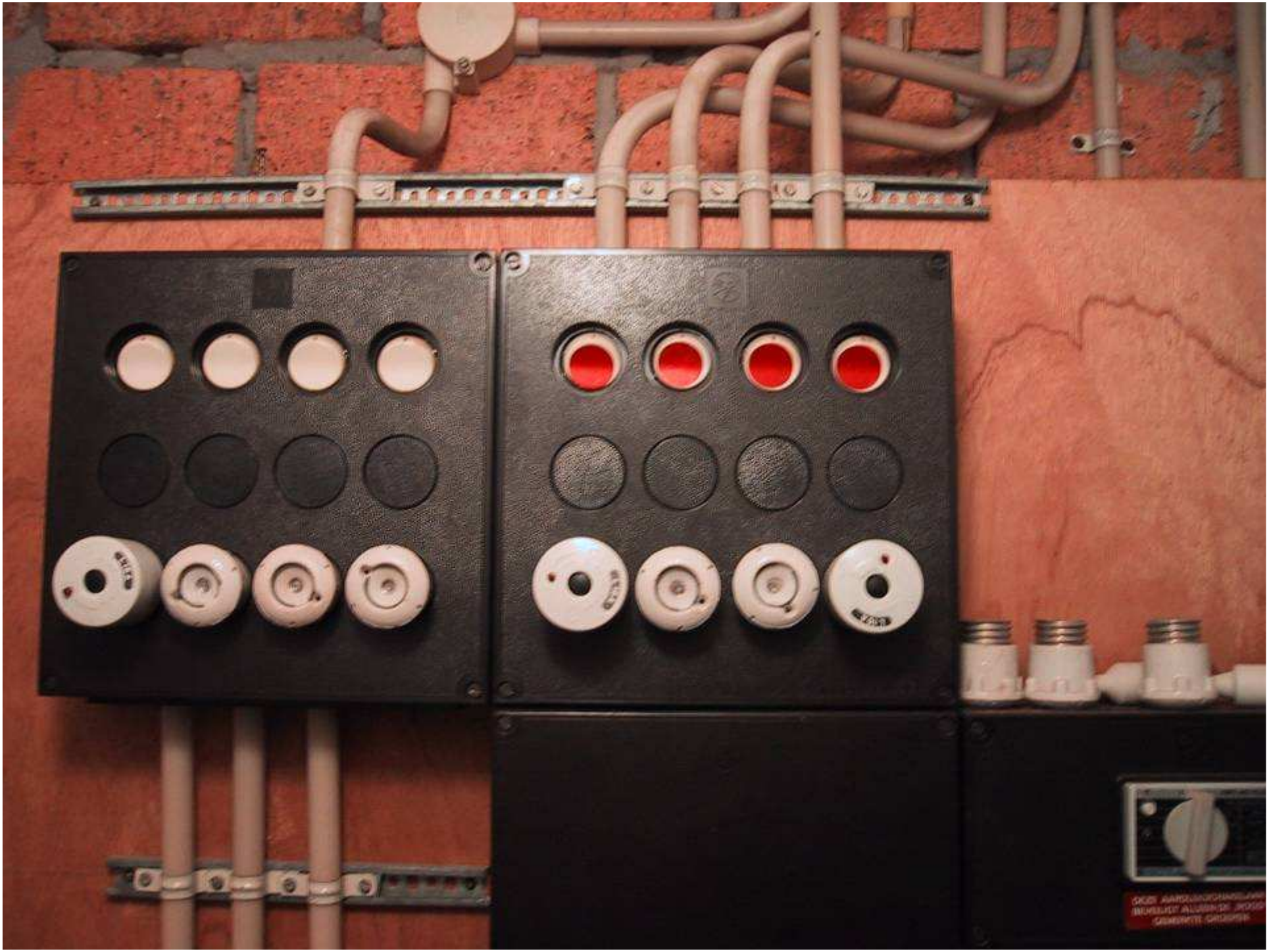
Targets and Biomarkers

CONNECTIVITY

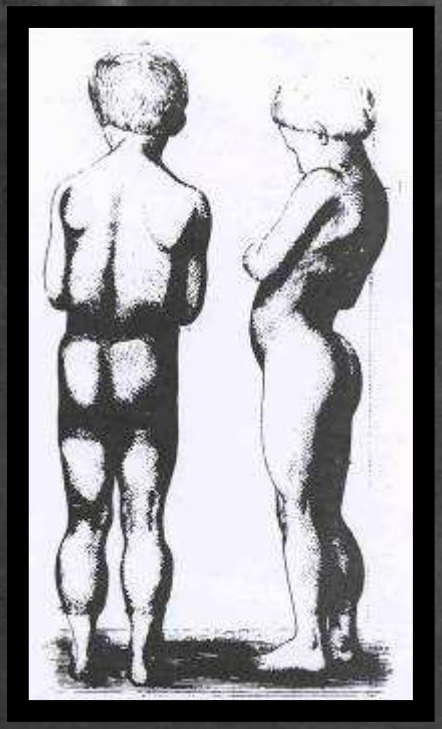
Maximalise impact of medical genomics:

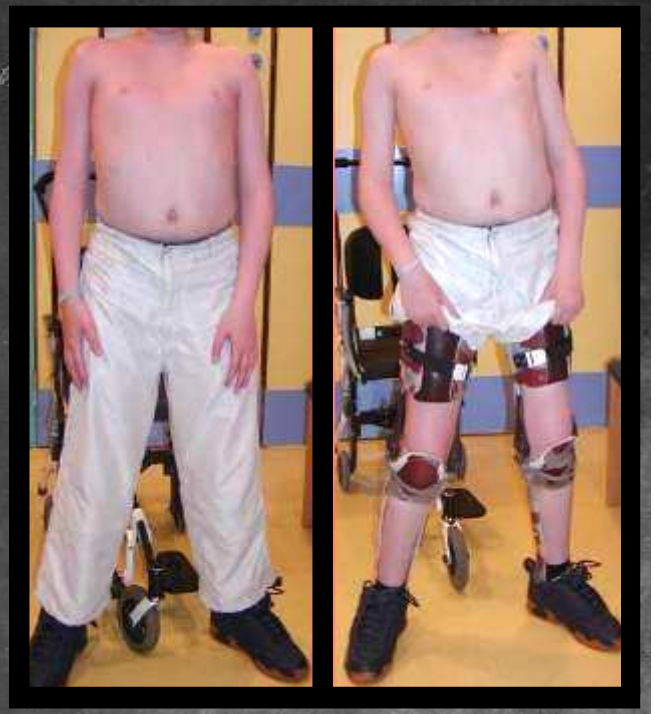
'Biomarker' discovery at system level





DUCHENNE MUSCULAR DYSTROPHY

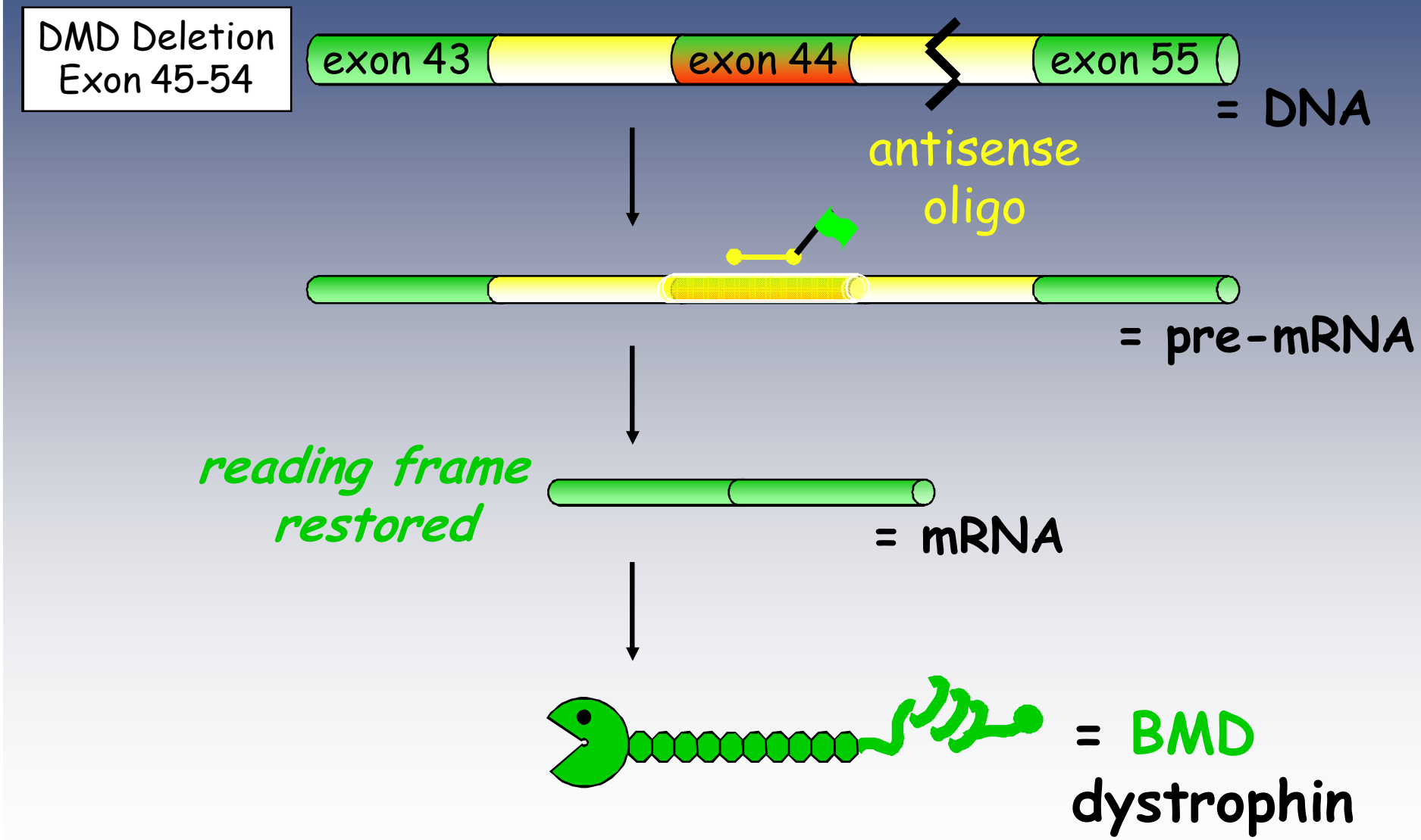




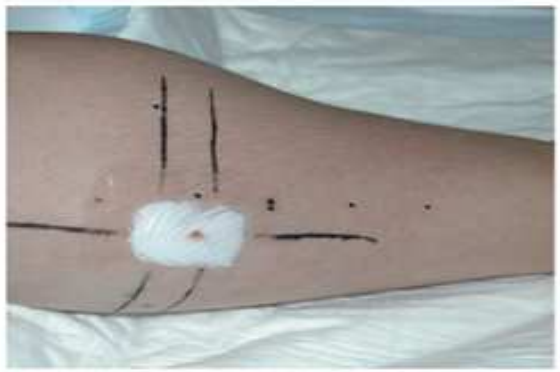
Most frequent lethal childhood disease

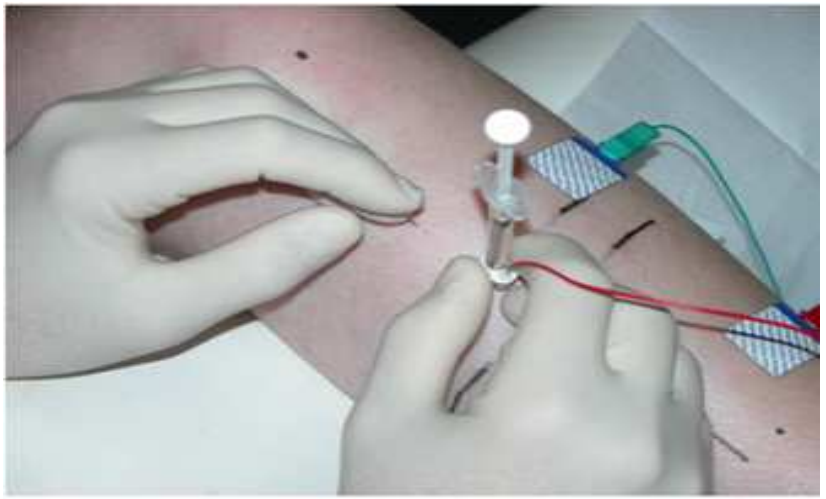
- **X-chromosomal, 1:3500 male newborns**
- **Defects in dystrophin gene (del/dup/mut)**
- **Largest gene in genome: 2.5 Mb, 79 exons**
  - **Out-of-frame deletions, duplications**
  - **Stop and splice mutations.**

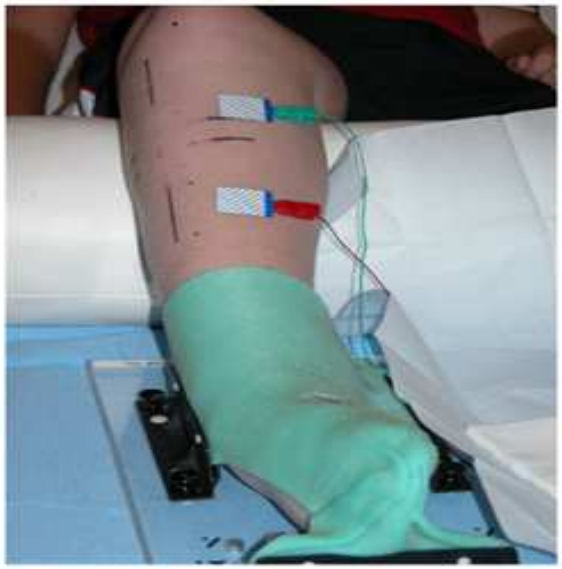
Mechanism: Antisense interference of splicing

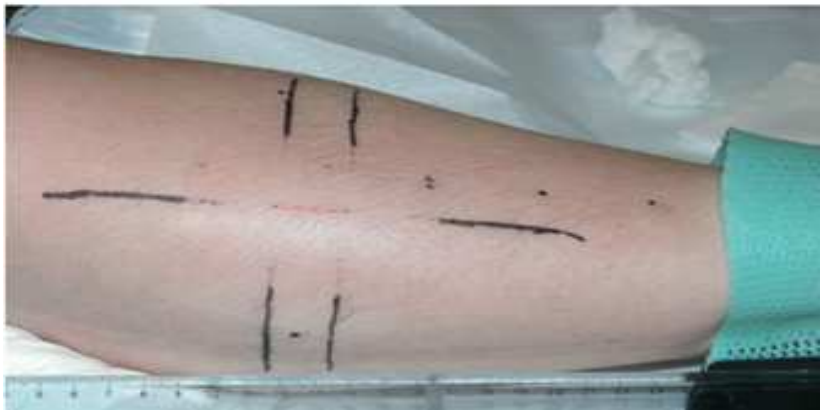


AON injection









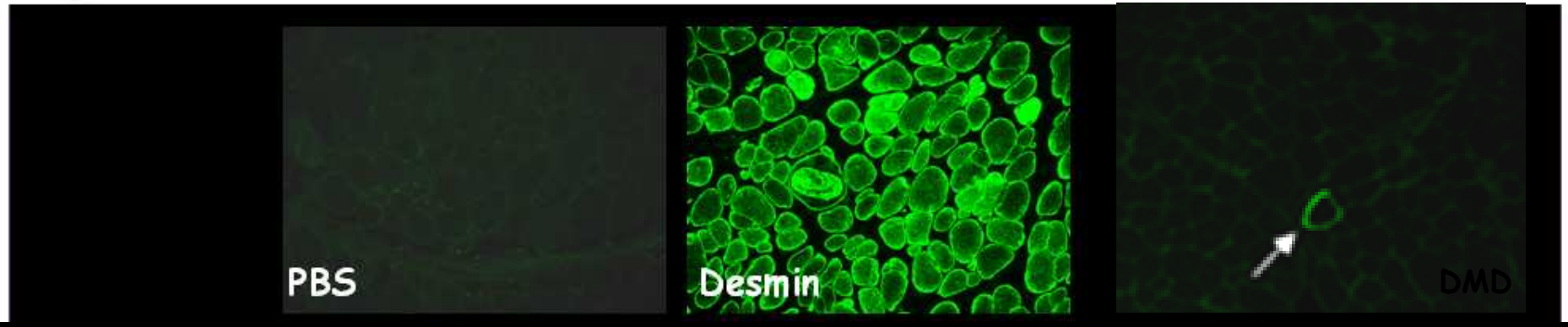
PROSENSA

EFFICACY ASSESSMENT Pt.1

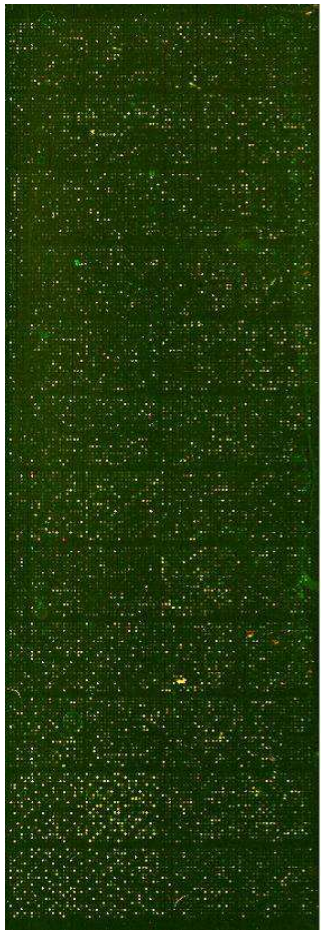
IMMUNOHISTOCHEMICAL ANALYSIS

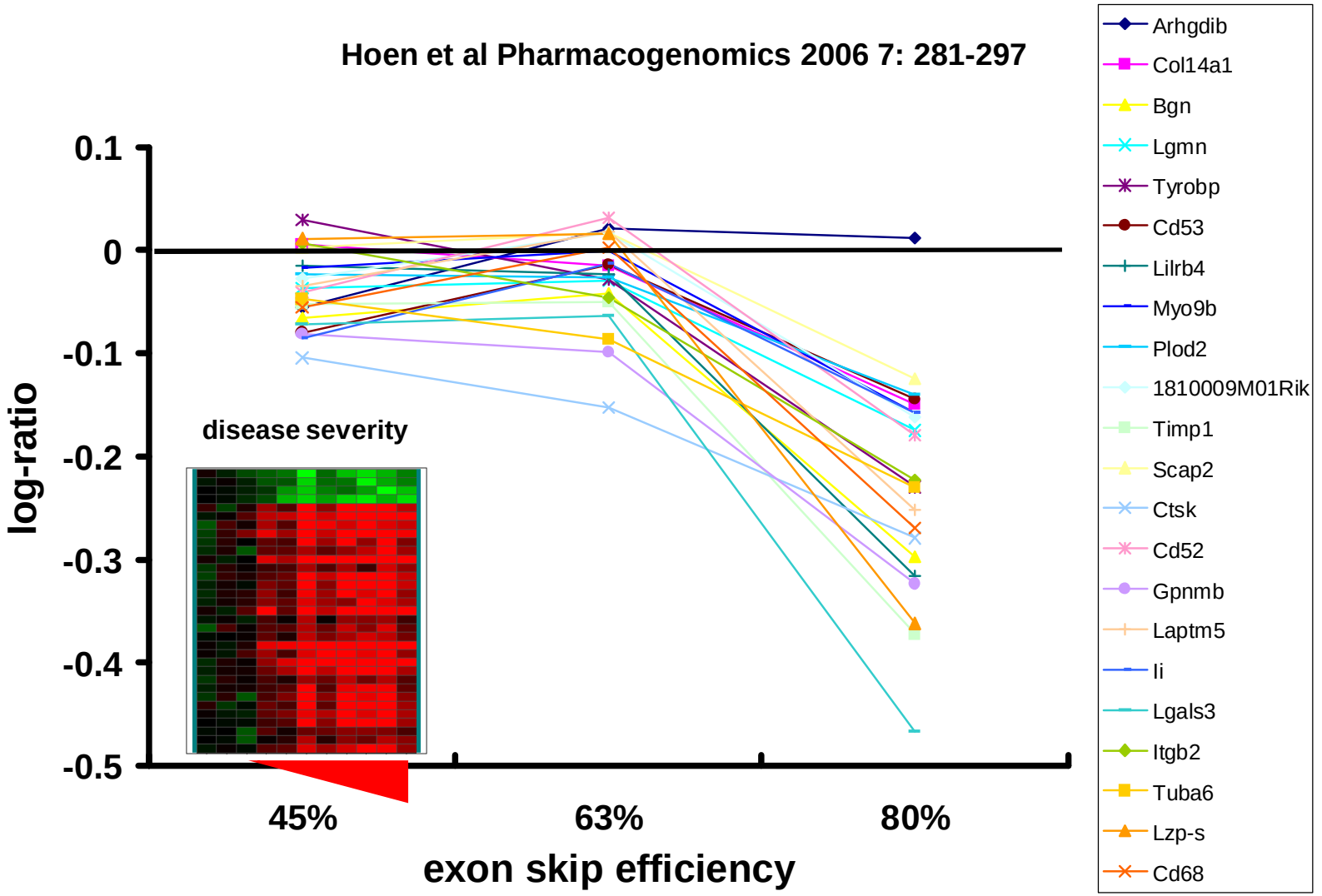
LUMC

van Deutekom et al. N Engl J Med. 357(26):2677-86 2007

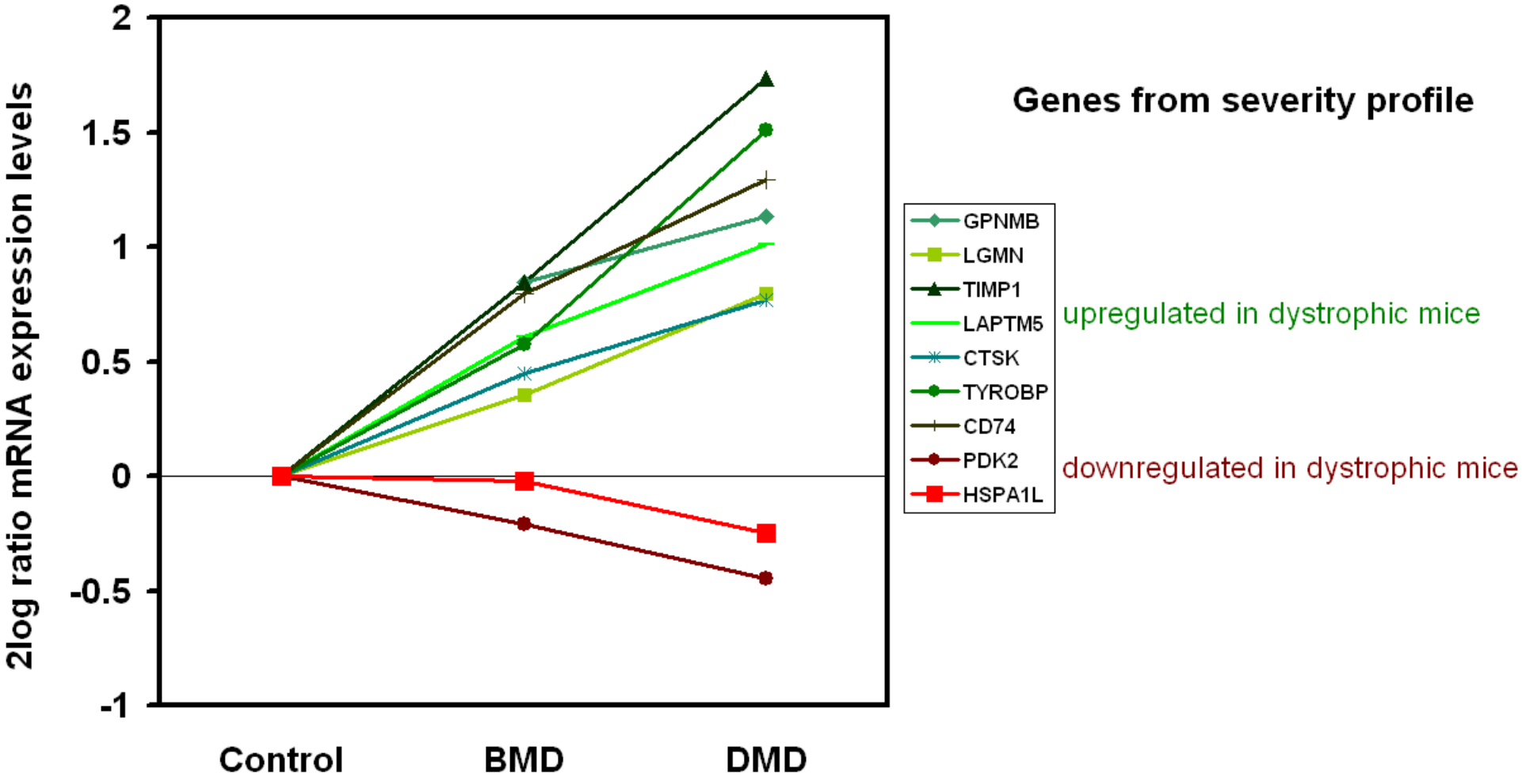


Genes upregulated in muscular dystrophy go down after rAAV-AON23 treatment





Translation to humans



Expression profiling[2]

- **model**
  - *wt / transgenic DCLK*
    - *constitutive expression δC-doublecortin-like kinase*
    - *brain > hippocampus*
  - *subtle behavioural abnormalities*
- **micro-array analysis 5 platforms**
  - *- only subtle changes*
  - *- biological replicates*

Pedotti et al BMC Genomics 2008, 9:124

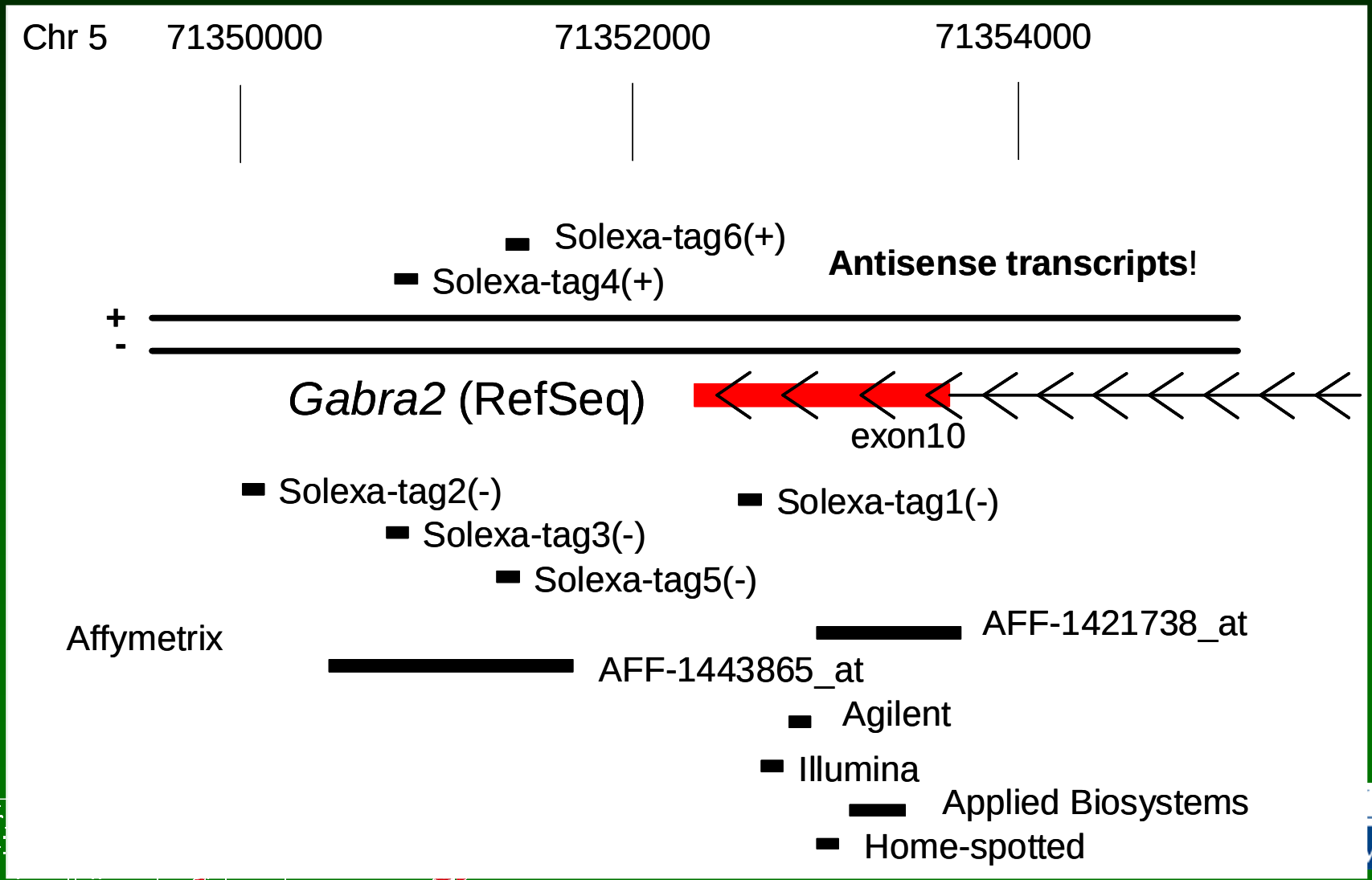
- **deep-sequence experiment**
  - *individual mice (Bio-replicates)*
    - *Leiden (n=5)*
  - *pools (wt / transgenics)*
    - *Illumina*

Hoen et al, NAR in press

Discovery properties

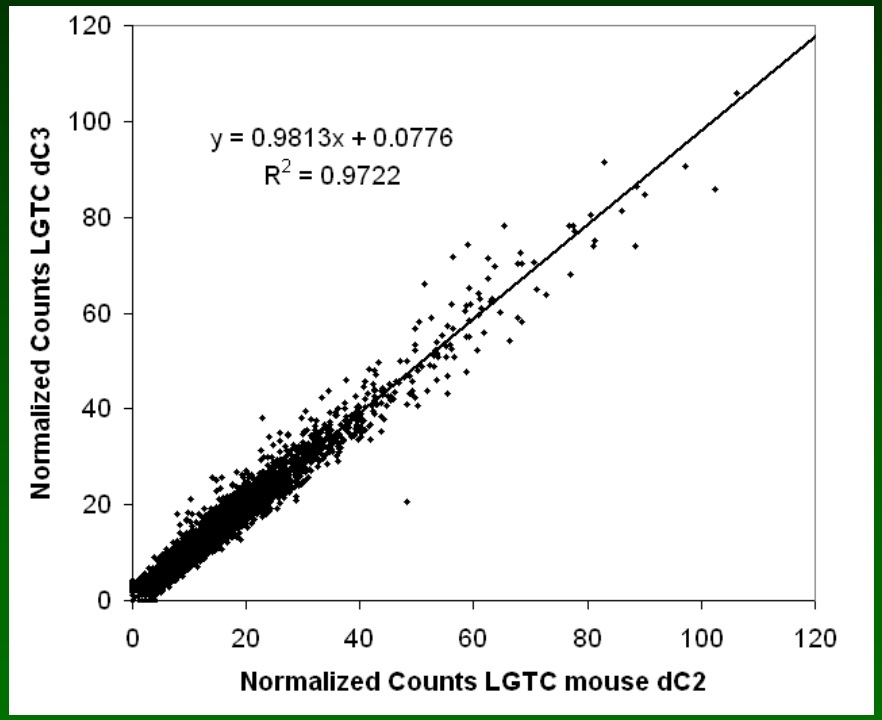
- 51% of genes in hippocampus show antisense transcription at > 2 transcripts per million (tpm)
- 47% of genes show alternative polyadenylation

Solexa finds 6 different transcripts for Gabra2



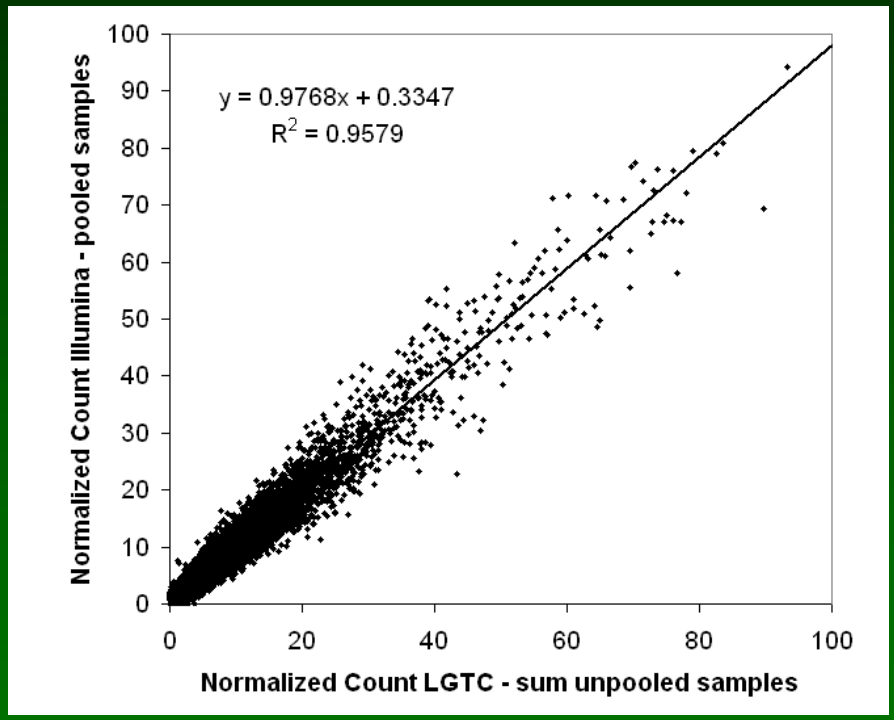
Lab2lab consistency

2 transgenic mice



(biological replicas)

2 different labs



(Illumina <> Leiden)

square root-transformed and scaled data

CENTRE FOR Medical Systems Biology

LUMC

Pooling

set differentially expressed genes

GENE	Name	Pool_WT	Pool_dC	WT1	WT3	WT4	WT6	dC1	dC2	dC3	dC4
Exosc8	Exosome component 8	14	0	28	2	0	0	0	0	1	0
Fgg	Fibrinogen, gamma polypeptide	60	0	72	1	1	0	0	0	0	0
Gc	Group specific component	22	0	41	0	1	0	0	0	0	0
Itih4	Inter alpha-trypsin inhibitor, heavy chain 4	26	0	51	1	1	0	0	1	0	0
Mug1	Murinoglobulin 1	20	0	25	0	0	0	0	0	0	0
Mup1	Major urinary protein 1	14	0	4	0	0	0	0	0	0	0
Mup1	Major urinary protein 1	18	0	8	0	0	0	0	0	0	0
Orm1	Orosomucoid 1	11	0	22	1	0	0	0	0	0	0
Rdh7	Retinol dehydrogenase 7	17	0	21	0	0	0	0	0	0	0
Serpina1a	Serine (or cysteine) peptidase inhibitor, clade A, member 1a	35	0	71	0	0	0	0	0	0	0
Serpina3k	Serine (or cysteine) peptidase inhibitor, clade A, member 3K	87	0	143	1	1	0	0	0	0	0

WT sample 2 contaminated with blood

Illumina deep sequencing compared with 5 microarrays

- Unbiased view of transcriptome
  not limited by array content
- Much deeper than corresponding SAGE
  ➔2 tpm vs. 91 tpm
- Antisense transcription detected
- Higher sensitivity: no ratio compression
  ➔fold change ratio's {
- Less data preprocessing:
  better, faster interpretation
- High inter-lab reproducibility
  no hyb/seq dependence

Illumina deep sequencing compared with 5 microarrays

- ➔Much easier to integrate deep sequencing data from different experiments and locations

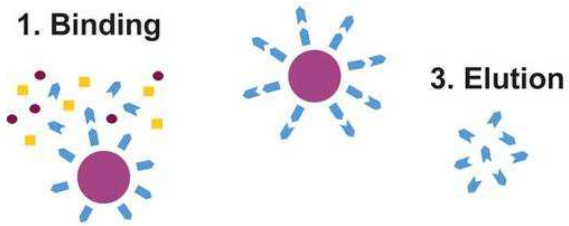
- ➔BIOBANKING applications!

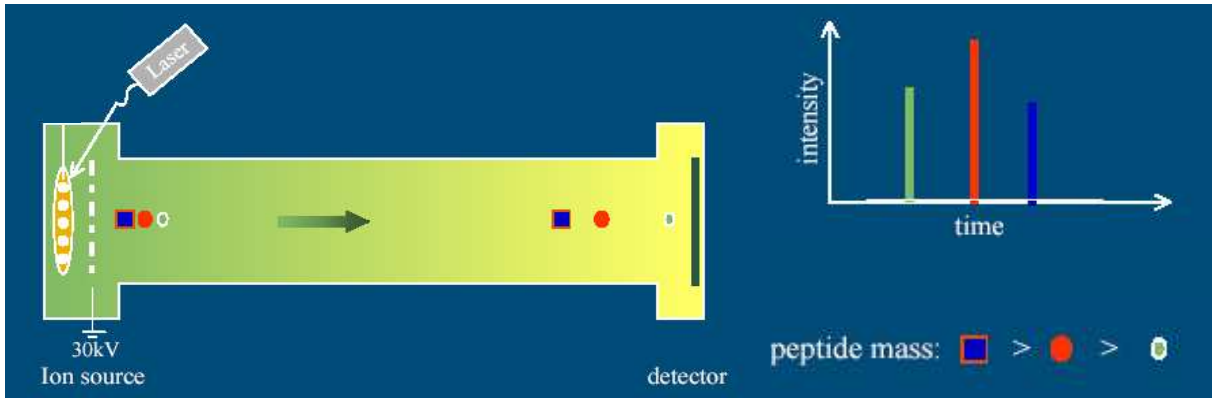
Serum protein profiling

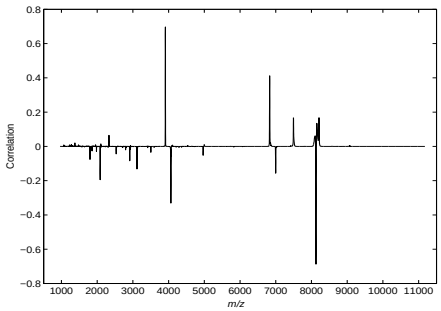
Muscle proteins leak into the blood: creatine kinase

Disease-specific profiles are expected

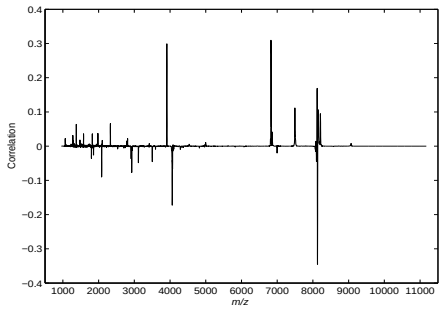
MALDI-TOF mass spectrometry



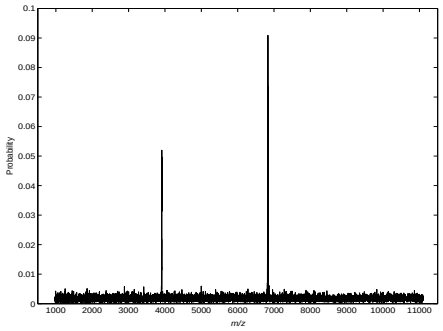




Lin



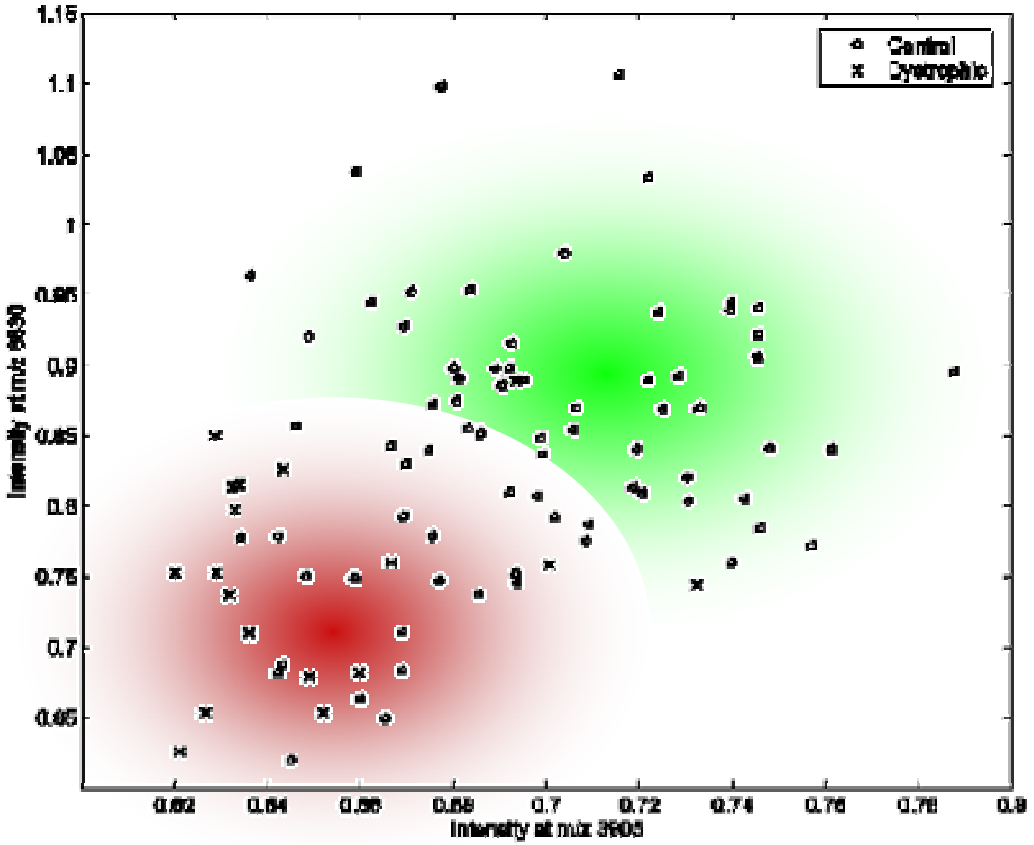
Log



Bayes

Mouse serum peptidomics 6

Fig. 4. Scatter plot of intensities of the *m/z* peaks 3905 against 6830 with class indicator.

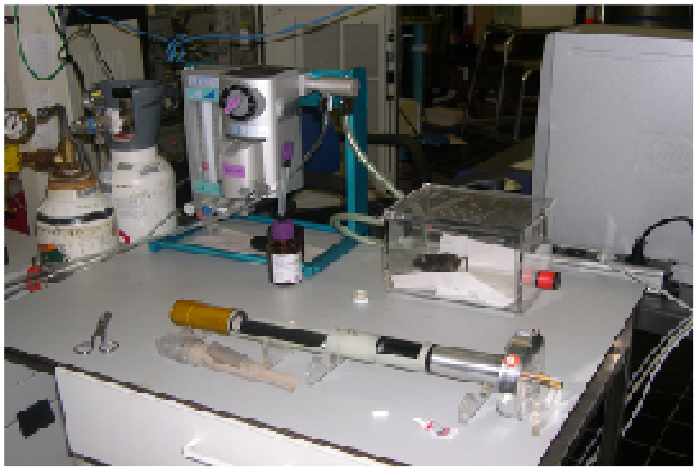


Magnetic Resonance Microscopy

- Fine anatomic detail in high resolution
- Functional aspects of motion
- Local metabolite quantitation

9.4 Tesla (400 MHz)

17.6 Tesla (750 MHz)

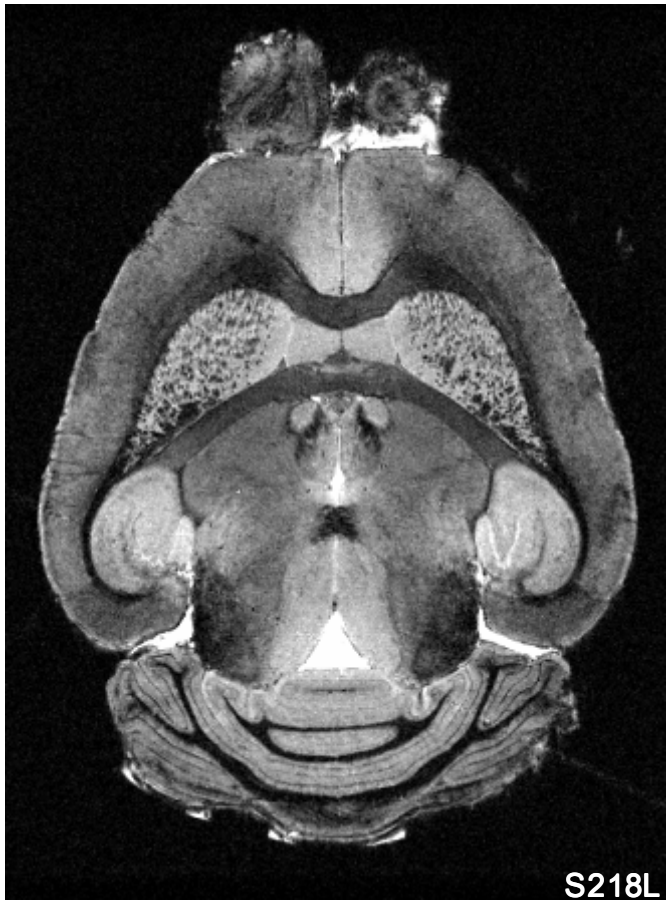


Anaesthesia

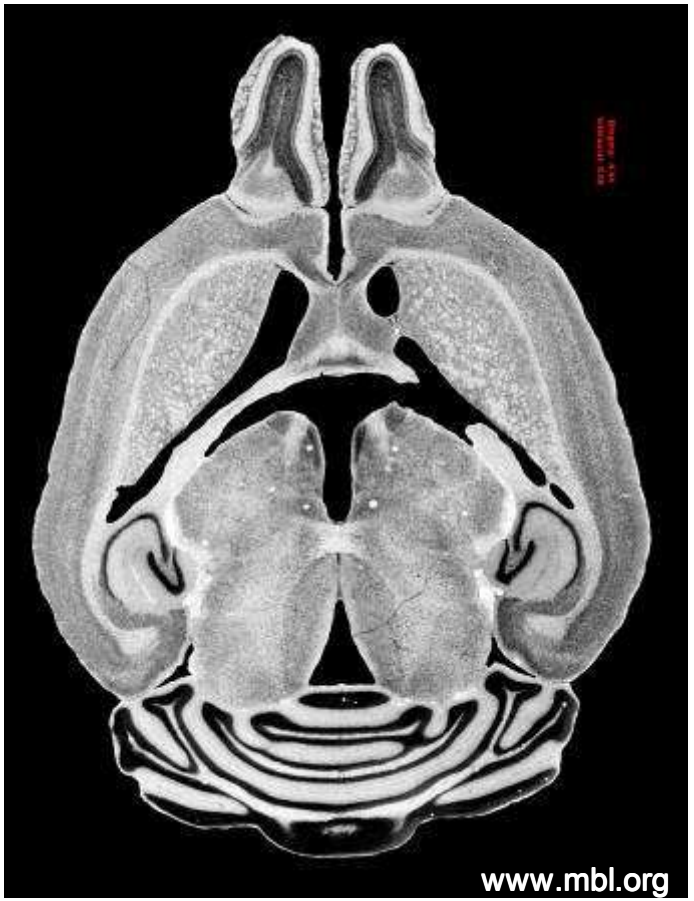
Mouse cavity

Prof. Rob Poelmann, LUMC
Prof. Huub de Groot, UL

High resolution *postmortem* imaging of S218L KI mouse brain



MRI



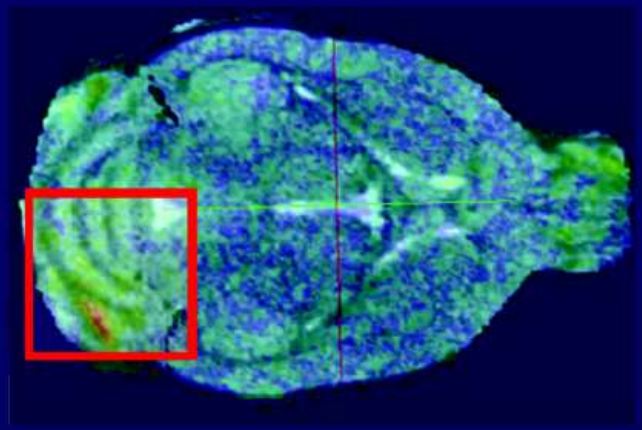
Histological Brain Atlas

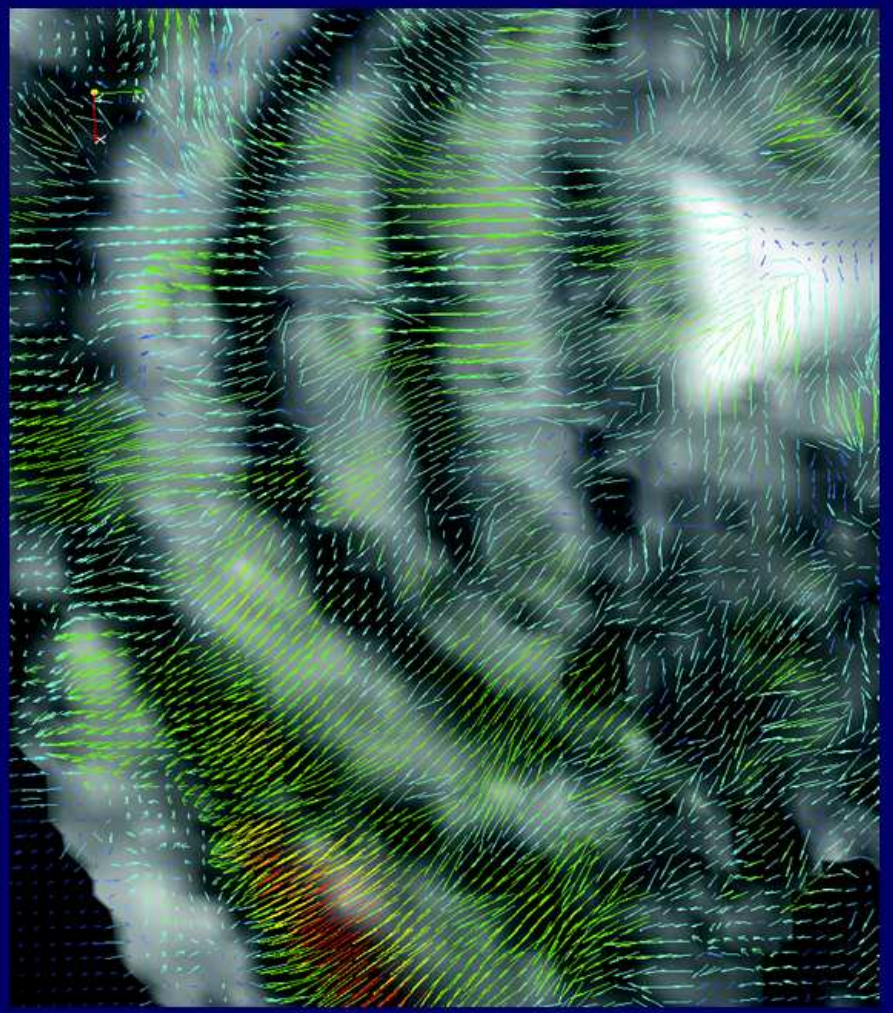
Postmortem; 9.4T; 30x30x30 µm; gadolinium contrast

Arn vd Maagdenberg/ R. Frants M.Ferrari LUMC
Louise van der Weerd, Bianca Hogers/ H de Groot, UL

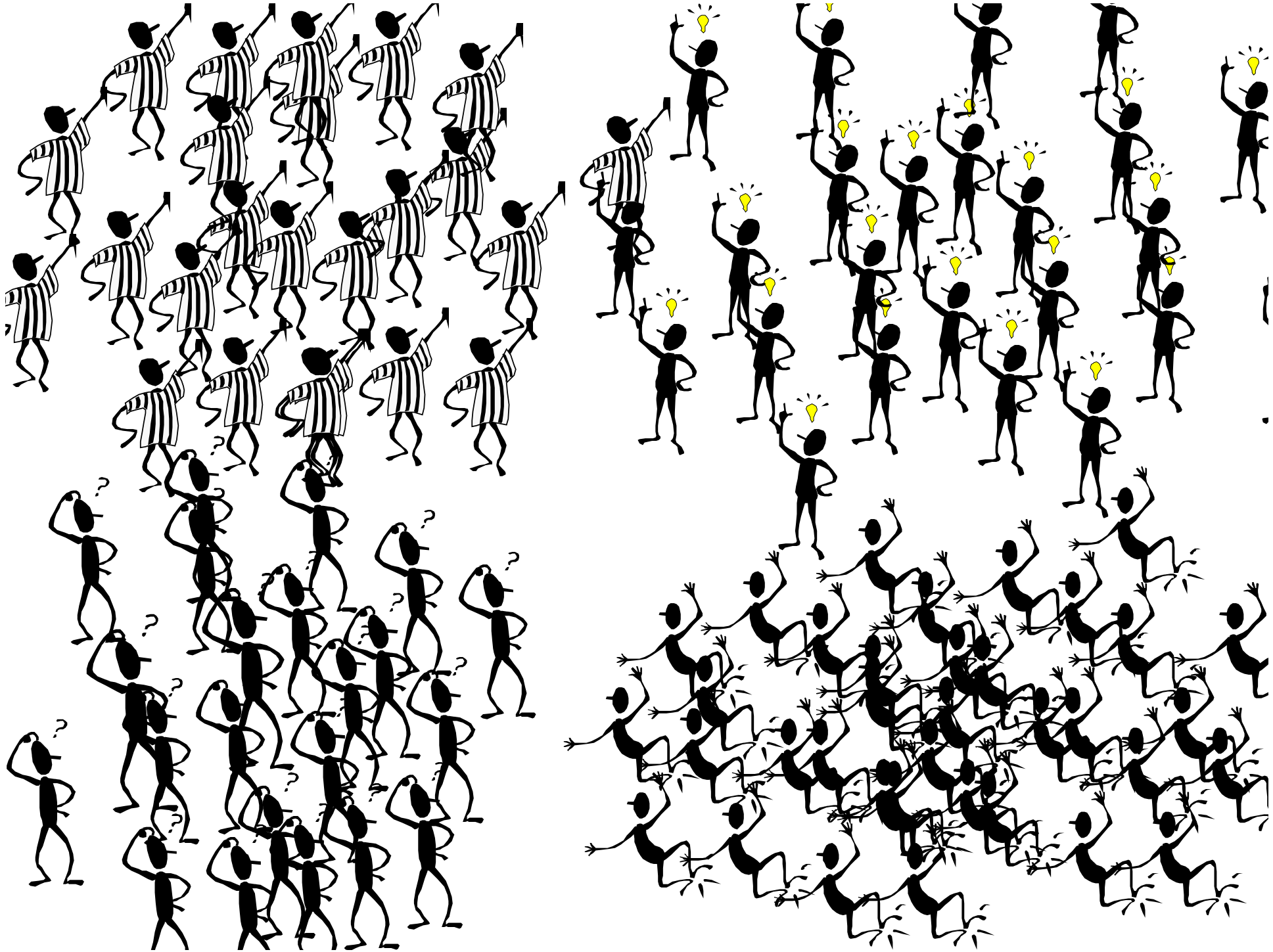
Functional and MRI studies

Deformation fields





Faiza Admiraal-Behloul,
Roald van der Laan



A High Proportion of Women in the General Population Carry the FGFR2 Breast Cancer Risk Allele

	Copies of FGFR2 Risk Allele		
	2	1	0
Frequency In UK Population	14%	47%	39%
Breast cancer Risk by age 70	10.5%	6.7%	5.5%

American Cancer Society Guidelines for Breast Screening with MRI as an Adjunct to Mammography

Debbie Saslow, PhD; Carla Boetes, MD, PhD; Wylie Burke, MD, PhD; Steven Harms, MD; Martin O. Leach, PhD; Constance D. Lehman, MD, PhD; Elizabeth Morris, MD; Etta Pisano,MD; Mitchell Schnall, MD, PhD; Stephen Sener, MD; Robert A. Smith, PhD; Ellen Warner,MD; Martin Yaffe, PhD; Kimberly S. Andrews; Christy A. Russell, MD (for the American Cancer Society Breast Cancer Advisory Group)

ABSTRACT New evidence on breast Magnetic Resonance Imaging (MRI) screening has become available since the American Cancer Society (ACS) last issued guidelines for the early detection of breast cancer in 2003. A guideline panel has reviewed this evidence and developed new recommendations for women at different defined levels of risk. **Screening MRI is recommended for women with an approximately 20–25% or greater lifetime risk of breast cancer**, including women with a strong family history of breast or ovarian cancer and women who were treated for Hodgkin disease. **There are several risk subgroups for which the available data are insufficient to recommend for or against screening, including women with a personal history of breast cancer, carcinoma in situ, atypical hyperplasia, and extremely dense breasts on mammography**. Diagnostic uses of MRI were not considered to be within the scope of this review.

CA Cancer J Clin 2007;57:75–89.

MRI Evaluation of the Contralateral Breast in Women with Recently Diagnosed Breast Cancer

Constance D. Lehman, M.D., Ph.D., Constantine Gatsonis, Ph.D., Christiane K. Kuhl, M.D., R. Edward Hendrick, Ph.D., Etta D. Pisano, M.D., Lucy Hanna, M.S., Sue Peacock, M.S., Stanley F. Smazal, M.D., Daniel D. Maki, M.D., Thomas B. Julian, M.D., Elizabeth R. DePeri, M.D., David A. Bluemke, M.D., Ph.D., and Mitchell D. Schnall, M.D., Ph.D., for the ACRIN Trial 6667 Investigators Group*

RESULTS

MRI detected clinically and mammographically occult breast cancer in the contralateral breast in 30 of 969 women who were enrolled in the study (3.1%). The sensitivity of MRI in the contralateral breast was 91%, and the specificity was 88%. The negative predictive value of MRI was 99%. A biopsy was performed on the basis of a positive MRI finding in 121 of the 969 women (12.5%), 30 of whom had specimens that were positive for cancer (24.8%); 18 of the 30 specimens were positive for invasive cancer. The mean diameter of the invasive tumors detected was 10.9 mm. The additional number of cancers detected was not influenced by breast density, menopausal status, or the histologic features of the primary tumor.

CONCLUSIONS

MRI can detect cancer in the contralateral breast that is missed by mammography and clinical examination at the time of the initial breast-cancer diagnosis. (ClinicalTrials.gov number, NCT00058058.)

Mammographic Density and the Risk and Detection of Breast Cancer

Norman F. Boyd, M.D., D.Sc., Helen Guo, M.Sc., Lisa J. Martin, Ph.D., Limei Sun, M.Sc., Jennifer Stone, M.Sc., Eve Fishell, M.D., F.R.C.P.C., Roberta A. Jong, M.D., F.R.C.P.C., Greg Hislop, M.D., F.R.C.P.C., Anna Chiarelli, Ph.D., Salomon Minkin, Ph.D., and Martin J. Yaffe, Ph.D.

CONCLUSIONS

Extensive mammographic density is strongly associated with the risk of breast cancer detected by screening or between screening tests. A substantial fraction of breast cancers can be attributed to this risk factor.

Future ahead

Soon we will be able to sequence a complete human genome,

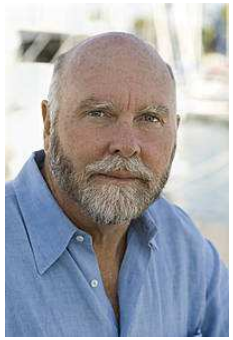
but if we can not make sense out of the variants detected,

as to whether they are "pathogenic or not",

this information is useless and prone to misinterpretation….

Human genomes

(individual genomes sequenced)



Craig Venter

James Watson

Marjolein Kriek

<u>ANONYMOUS:</u>
Yoruban HapMap male
Yoruban HapMap trio
Asian genome
....

A human genome …

(www.LUMC.nl)

- **by academic hospital**

  not a large genome center,
  nor a company (sequence technology)

- **Marjolein Kriek**

  PhD, clinical geneticist (i.t.)

  first from LUMC,
  Leiden,
  Nederland,
  Europe
  female world-wide

CENTRE FOR
Medical Systems Biology

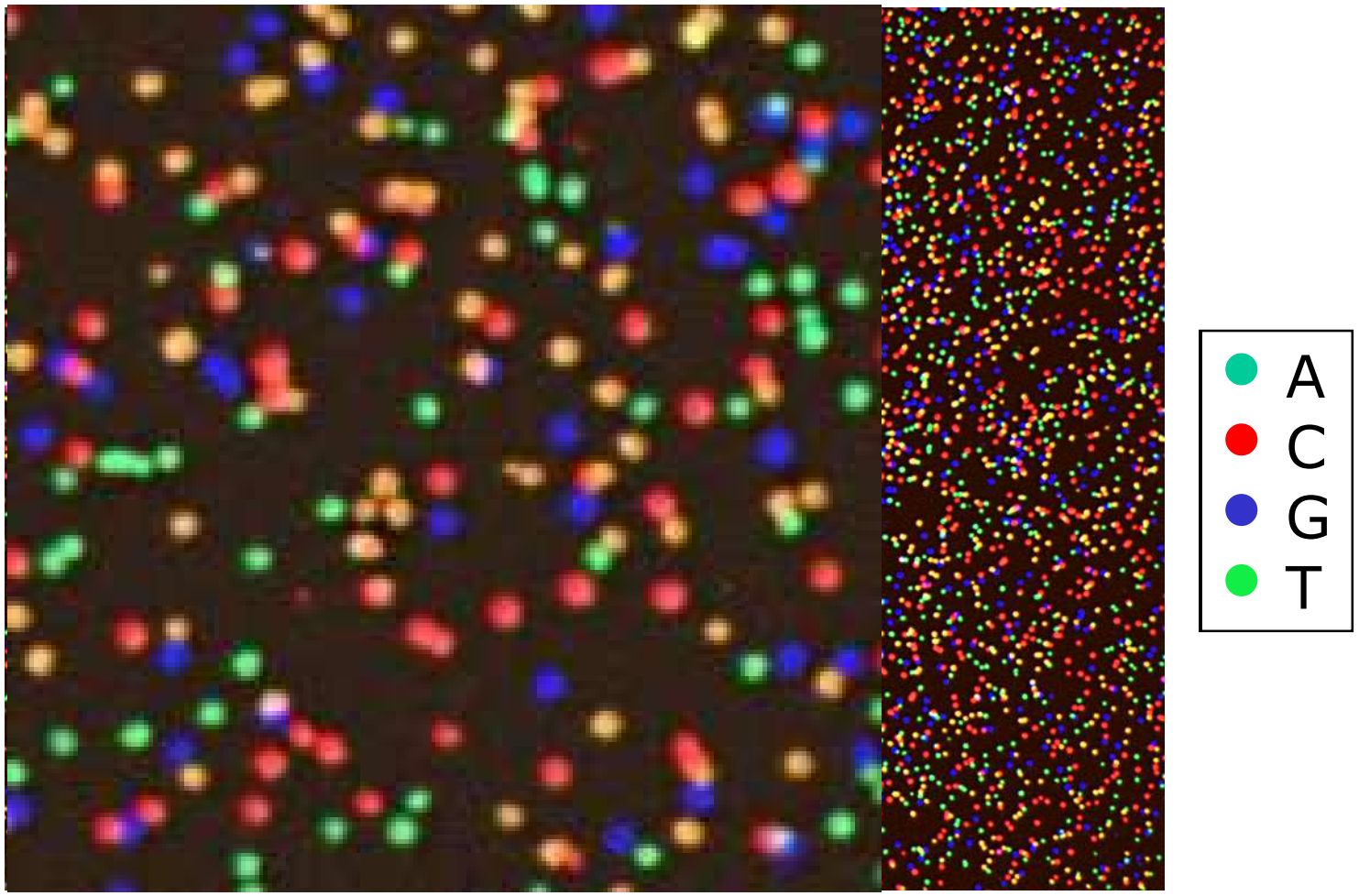
A human genome …

- **why us?**
  - *show it is possible*
    - *technical, computational, analytical*
  - *to learn*
    - *technology, data floods, analysis*
  - *attractive project to tackle*
- **why her ?**
  - *clinical geneticist*
  - *X-chromosome less variable*
    - *look at more, not fewer*
- **results**
  - *technically - no problem*
  - *computationally - at our limits*
  - *analytically - not (yet) possible*
  - *- as expected*
- **>> to be applied in patients**
  - *resolve cause genetic disease*

CENTRE FOR Medical Systems Biology

LUMC

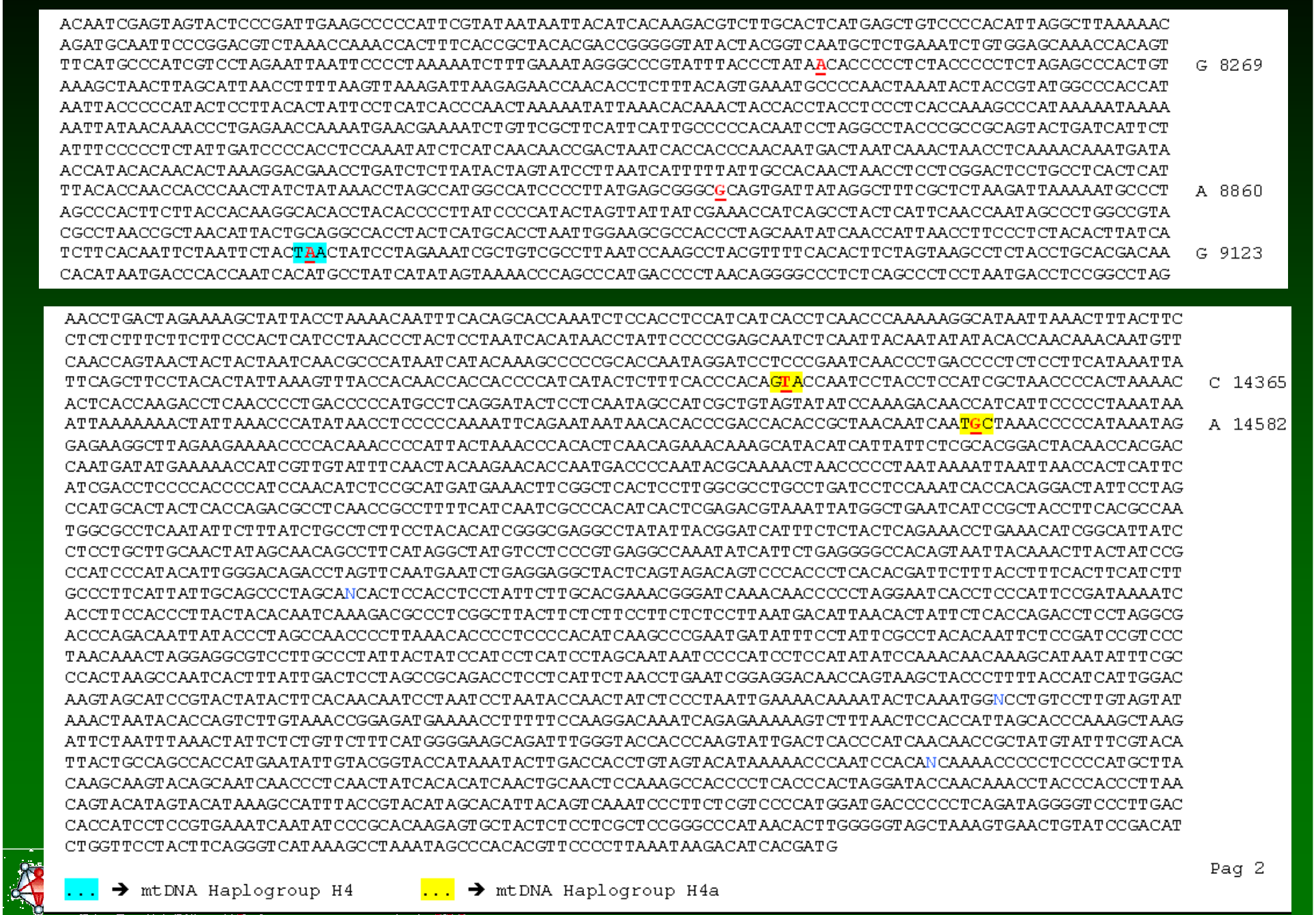
First cycle



Analysis

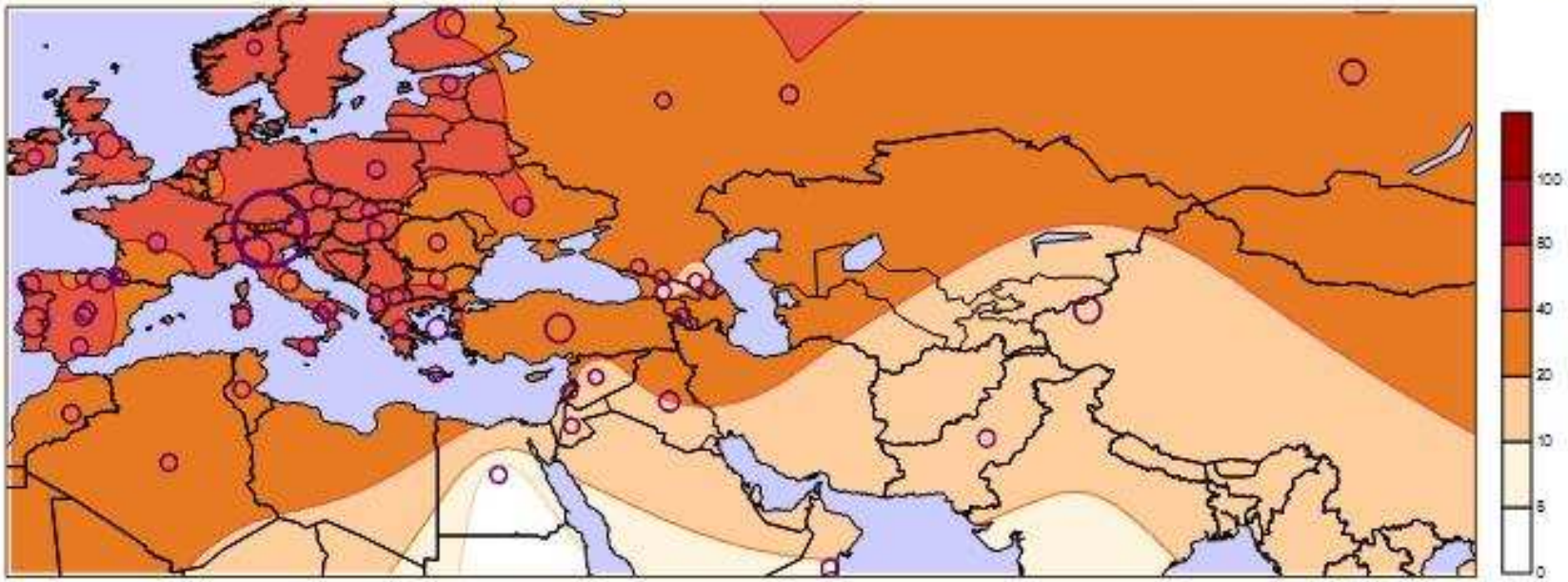
draw DNA-based conclusions

1. a female (no Y-chromosome sequences)

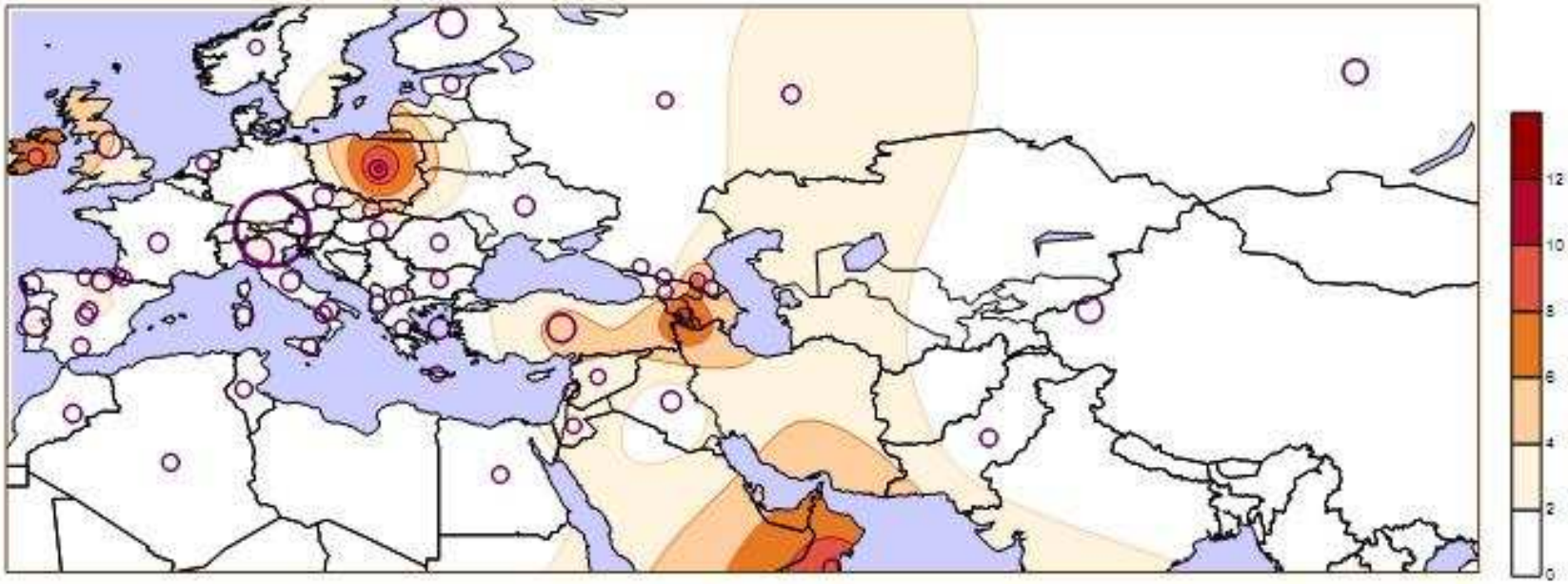

ACAATCGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACGTCTTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAAC
AGATGCAATTCCCGGACGTCTAAACCAAACCACTTTCACCGCTACACGACCGGGGGTATACTACGGTCAATGCTCTGAAATCTGTGGAGCAAACCACAGT
TTCATGCCCATCGTCCTAGAATTAATTCCCCTAAAAATCTTTGAAATAGGGCCCGTATTTACCCTATAACACCCCCTCTACCCCCTCTAGAGCCCACTGT
G 8269
AAAGCTAACTTAGCATTAACCTTTTAAGTTAAAGATTAAGAGAACCAACACCTCTTTACAGTGAAATGCCCCAACTAAATACTACCGTATGGCCCACCAT
AATTACCCCCATACTCCTTACACTATTCCTCATCACCCAACTAAAAATATTAAACACAAACTACCACCTACCTCCCTCACCAAAGCCCATAAAAATAAAA
AATTATAACAAACCCTGAGAACCAAAATGAACGAAAATCTGTTCGCTTCATTCATTGCCCCCACAATCCTAGGCCTACCCGCCGCAGTACTGATCATTCT
ATTTCCCCCTCTATTGATCCCCACCTCCAAATATCTCATCAACAACCGACTAATCACCACCCAACAATGACTAATCAAACTAACCTCAAAACAAATGATA
ACCATACACAACACTAAAGGACGAACCTGATCTCTTATACTAGTATCCTTAATCATTTTTATTGCCACAACTAACCTCCTCGGACTCCTGCCTCACTCAT
TTACACCAACCACCCAACTATCTATAAACCTAGCCATGGCCATCCCCTTATGAGCGGGCGCAGTGATTATAGGCTTTCGCTCTAAGATTAAAAATGCCCT
A 8860
AGCCCACTTCTTACCACAAGGCACACCTACACCCCTTATCCCCATACTAGTTATTATCGAAACCATCAGCCTACTCATTCAACCAATAGCCCTGGCCGTA
CGCCTAACCGCTAACATTACTGCAGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCCTAGCAATATCAACCATTAACCTTCCCTCTACACTTATCA
TCTTCACAATTCTAATTCTACTAACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACTTCTAGTAAGCCTCTACCTGCACGACAA
G 9123
CACATAATGACCCACCAATCACATGCCTATCATATAGTAAAACCCAGCCCATGACCCCTAACAGGGGCCCTCTCAGCCCTCCTAATGACCTCCGGCCTAG
AACCTGACTAGAAAAGCTATTACCTAAAACAATTTCACAGCACCAAATCTCCACCTCCATCATCACCTCAACCCAAAAAGGCATAATTAAACTTTACTTC
CTCTCTTTCTTCTTCCCACTCATCCTAACCCTACTCCTAATCACATAACCTATTCCCCCGAGCAATCTCAATTACAATATATACACCAACAAACAATGTT
CAACCAGTAACTACTACTAATCAACGCCCATAATCATACAAAGCCCCCGCACCAATAGGATCCTCCCGAATCAACCCTGACCCCTCTCCTTCATAAATTA
TTCAGCTTCCTACACTATTAAAGTTTACCACAACCACCACCCCATCATACTCTTTCACCCACAGTACCAATCCTACCTCCATCGCTAACCCCACTAAAAC
C 14365
ACTCACCAAGACCTCAACCCCTGACCCCCATGCCTCAGGATACTCCTCAATAGCCATCGCTGTAGTATATCCAAAGACAACCATCATTCCCCCTAAATAA
ATTAAAAAAAACTATTAAACCCATATAACCTCCCCCAAAATTCAGAATAATAACACACCCGACCACACCGCTAACAATCAATGCTAAACCCCCATAAATAG
A 14582
GAGAAGGCTTAGAAGAAAACCCCACAAACCCCCATTACTAAACCCACACTCAACAGAAACAAAGCATACATCATTATTCTCGCACGGACTACAACCACGAC
CAATGATATGAAAAACCATCGTTGTATTTCAACTACAAGAACACCAATGACCCCAATACGCAAAACTAACCCCCTAATAAAATTAATTAACCACTCATTC
ATCGACCTCCCCACCCCATCCAACATCTCCGCATGATGAAACTTCGGCTCACTCCTTGGCGCCTGCCTGATCCTCCAAATCACCACAGGACTATTCCTAG
CCATGCACTACTCACCAGACGCCTCAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACGTAAATTATGGCTGAATCATCCGCTACCTTCACGCCAA
TGGCGCCTCAATATTCTTTATCTGCCTCTTCCTACACATCGGGCGAGGCCTATATTACGGATCATTTCTCTACTCAGAAACCTGAAACATCGGCATTATC
CTCCTGCTTGCAACTATAGCAACAGCCTTCATAGGCTATGTCCTCCCGTGAGGCCAAATATCATTCTGAGGGGCCACAGTAATTACAAACTTACTATCCG
CCATCCCATACATTGGGACAGACCTAGTTCAATGAATCTGAGGAGGCTACTCAGTAGACAGTCCCACCCTCACACGATTCTTTACCTTTCACTTCATCTT
GCCCTTCATTATTGCAGCCCTAGCANCACTCCACCTCCTATTCTTGCACGAAACGGGATCAAACAACCCCCTAGGAATCACCTCCCATTCCGATAAAATC
ACCTTCCACCCTTACTACACAATCAAAGACGCCCTCGGCTTACTTCTCTTCCTTCTCTCCTTAATGACATTAACACTATTCTCACCAGACCTCCTAGGCG
ACCCAGACAATTATACCCTAGCCAACCCCTTAAACACCCCTCCCCACATCAAGCCCGAATGATATTTCCTATTCGCCTACACAATTCTCCGATCCGTCCC
TAACAAACTAGGAGGCGTCCTTGCCCTATTACTATCCATCCTCATCCTAGCAATAATCCCCATCCTCCATATATCCAAACAACAAAGCATAATATTTCGC
CCACTAAGCCAATCACTTTATTGACTCCTAGCCGCAGACCTCCTCATTCTAACCTGAATCGGAGGACAACCAGTAAGCTACCCTTTTACCATCATTGGAC
AAGTAGCATCCGTACTATACTTCACAACAATCCTAATCCTAATACCAACTATCTCCCTAATTGAAAACAAAATACTCAAATGGNCCTGTCCTTGTAGTAT
AAACTAATACACCAGTCTTGTAAACCGGAGATGAAAACCTTTTTCCAAGGACAAATCAGAGAAAAAGTCTTTAACTCCACCATTAGCACCCAAAGCTAAG
ATTCTAATTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACA
TTACTGCCAGCCACCATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCCAATCCACANCAAAACCCCCTCCCCATGCTTA
CAAGCAAGTACAGCAATCAACCCTCAACTATCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATACCAACAAACCTACCCACCCTTAA
CAGTACATAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCAAATCCCTTCTCGTCCCCATGGATGACCCCCCTCAGATAGGGGTCCCTTGAC
CACCATCCTCCGTGAAATCAATATCCCGCACAAGAGTGCTACTCTCCTCGCTCCGGGCCCATAACACTTGGGGGTAGCTAAAGTGAACTGTATCCGACAT
CTGGTTCCTACTTCAGGGTCATAAAGCCTAAATAGCCCACACGTTCCCCTTAAATAAGACATCACGATG
Pag 2
... → mtDNA Haplogroup H4
... → mtDNA Haplogroup H4a

Distribution of mtDNA Haplogroup H

(source of data www.genebase.com)



Distribution of mtDNA Haplogroup H4



BBMRI Benefits - 1. Health

A more precise (biology-based ?) classification of disease will

speed up the development of more effective (and cost-effective) treatment

reduce undesired side effects of medicines,

improve success in clinical trial design

➔ lead to new concepts of disease prevention and health promotion

Benefits – 2. Socio-economic

Chronic and slowly progressive complex diseases cause a large economic burden across Europe.

Population Survey and Biobanking research will:

- **lead to improvements in disease prevention & treatment**
- **increase healthy years and quality of life (across European Union and outside)**
- **reduce the need for health care resources (refocusing)**
- **increase work capacity of the European population**
- **immense positive economic impact**

➔ justifies the large investments required to establish and maintain European biobanking infrastructure

Benefits – 3. Industry

Life sciences and biotechnology widely regarded as most promising frontier technology coming decade.

Pan-European Biobanking will:

- **stimulate research activities across European countries**
- **foster new synergies between the industrial and academic sectors**
- **strengthen the competitiveness of European biotech and health-related industry**

Benefits – 3. Industry

Final goal:

improved prevention and therapy,

Short-term benefit:

development of much more powerful diagnostic tools.

Molecular diagnostics (exploiting 'omics' to classify disease and to identify individuals at risk): *fastest growing segment in the health care industry*

Most Critical Issues

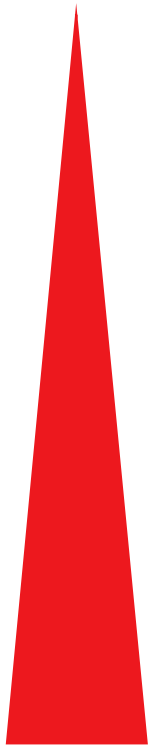
Standardization

Regulation of access

Incentives for contributors

Legal and ethical constrains

Sustained funding



Low risk

High risk

Ethical, legal, societal (political)

Challenges

- Optimizing resources and sustaining the high level of health care in Europe
- Changing economical and environmental landscape of the 21st century
- Crisis- and misunderstanding-ridden debates on food safety, cloning and gene patenting
- Mosaic and complex legal framework - while starting points are quite //

Ethical, legal, societal (political) -2

➔ **Improving cross-communication between societal disciplines - aligning the regulatory frameworks - is a prime opportunity to improve communication and foster better understanding:**

- **between the scientific, medical, legislative and social disciplines**
- **between the professionals, the patient communities and the public at large**

➔ **Life sciences and biotechnology, major fields of advances, will have a central role in the knowledge society of the near future**

Coordination

Coordinate the field:

GenomeEUtwin (FP5)

POPULATION BIOBANKS, EUHEALTHGEN, COGENE and PHOEBE (FP6)

ENGAGE, USING BIOBANKS (FP7)

ESFRI Roadmap: BBMRI: Biobanking and Molecular Repositories Infrastructure

P3G: The Public Population Project on Genomics: Global movement across international boundaries.

Harmonization of standards and procedures should be at a global level i.e. the efforts of P3G and PHOEBE should not be duplicated but built upon.

Coordination

But:

Coordinate the coordinators, too!

- **EU,**
- **ESFRI,**
- **ESF,**
- **OECD,**
- **UNESCO,**
- **WHO**
- **National agencies**