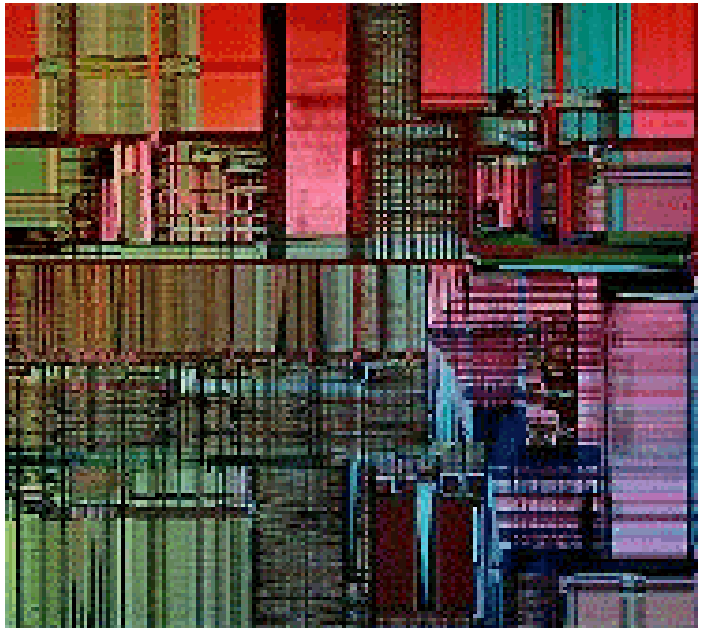
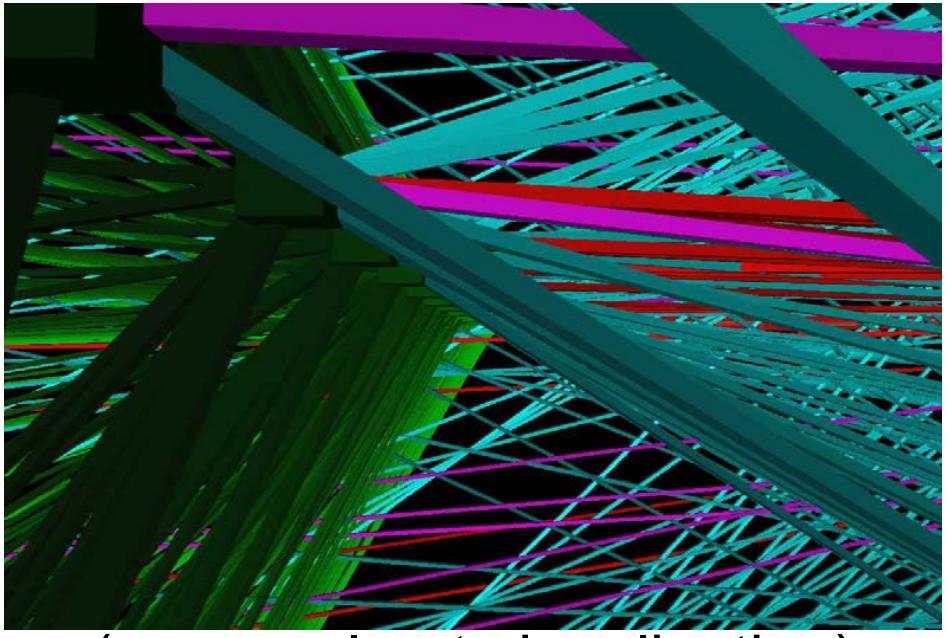
Life is the translation of the information in the genome into the phenotype of the organism:

The organism ‚computes‘ this phenotype from its genotype, given a specific environment

? Genome !

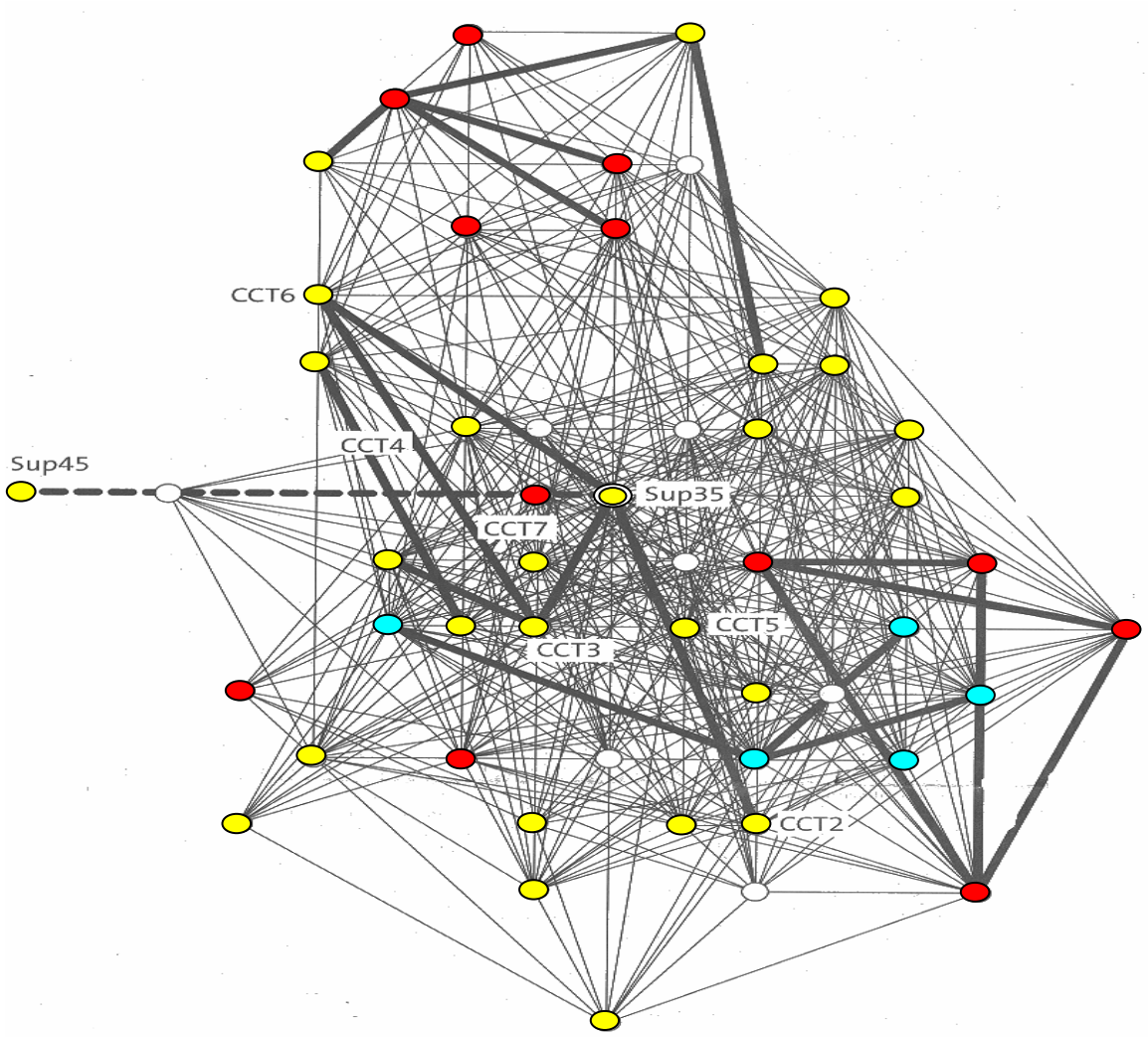


(PentiumV)



(neuronal net visualisation)

Phenotype


CCT6
CCT4
Sup45
Sup35
CCT7
CCT5
CCT3
CCT2

The Problem

World wide ~12 million new cancer cases/year

Cure rates for many common forms of cancer have hardly changed over the last decades

Even the most advanced targeted therapies are typically only effective for a small fraction of the patients

Pharma development costs have dramatically increased, while the number of new drugs keeps dropping

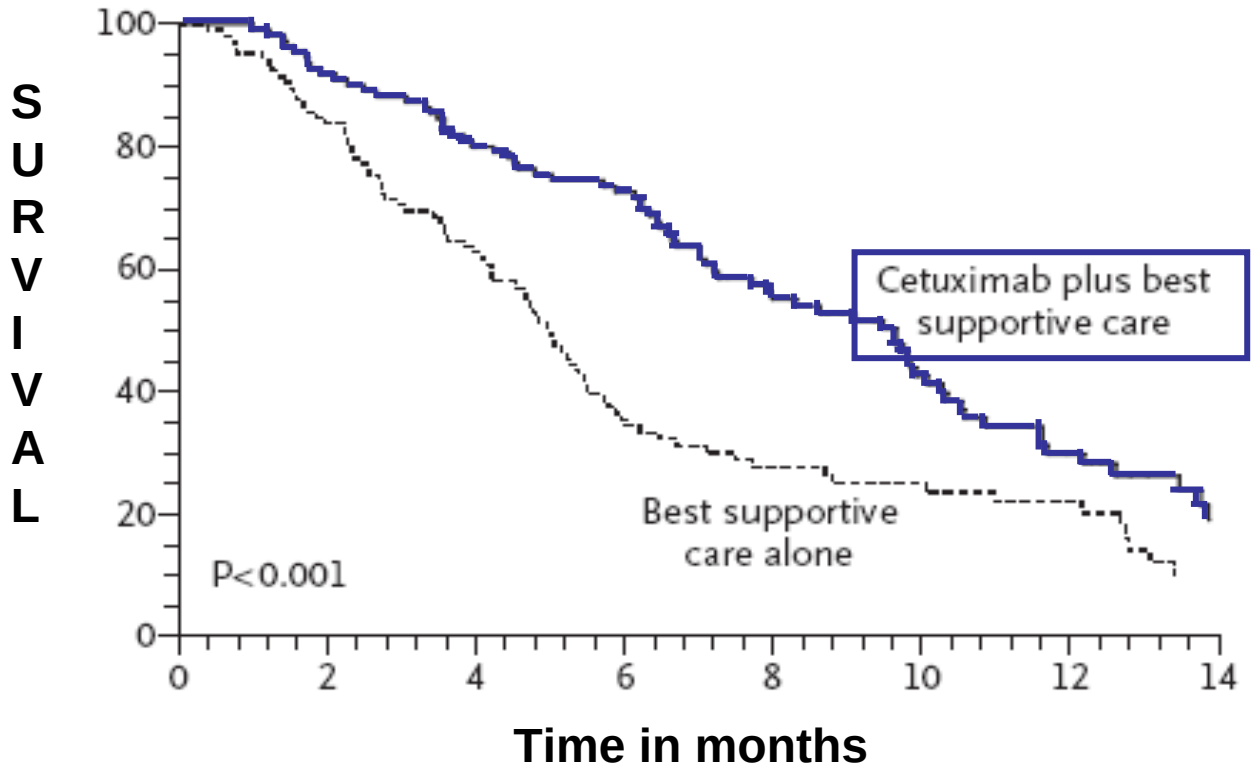
Global Cancer Burden (millions)

	2008	2030
New cases	12.5	26.4
Deaths	7.6	17.1
Alive with cancer	28.0	80.0

Boyle and Levin, 2009

Colon Cancer

(K-ras wild type)

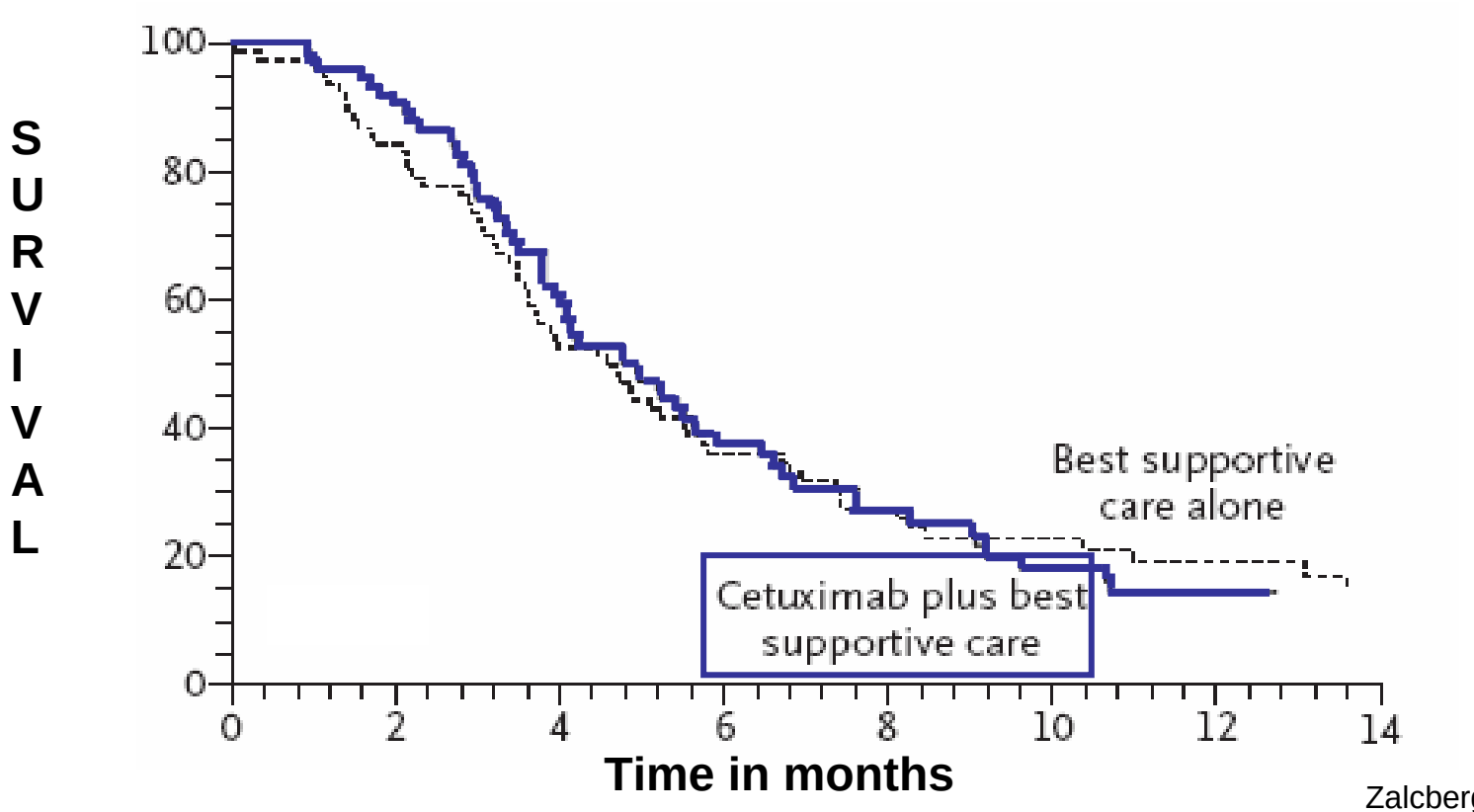


Zalcberg et al, NEJM 2008

Many colon cancer patients with tumors without K-ras wild type will respond positively to treatment with EGFR receptor anatagonists

Colon Cancer

Mutated K-ras



Zalcberg et al, NEJM 2008

Patients with tumors with K-ras mutations do not: a 30000 Euro treatment will show no effect on the tumor (but will cause significant side effects on the patient)

Der Anfang: 1973

Proc. Nat. Acad. Sci. USA
Vol. 70, No. 12, Part I, pp. 3581–3584, December 1973

The Nucleotide Sequence of the *lac* Operator

(regulation/protein–nucleic acid interaction/DNA–RNA sequencing/oligonucleotide priming)

WALTER GILBERT AND ALLAN MAXAM

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Communicated by J. D. Watson, August 9, 1973

ABSTRACT **The *lac* repressor protects the *lac* operator against digestion with deoxyribonuclease. The protected fragment is double-stranded and about 27 base-pairs long. We determined the sequence of RNA transcription copies of this fragment and present a sequence for 24 base pairs. It is:**

5′--T G G A A T T G T G A G C G G A T A A C A A T T 3′
3′--A C C T T A A C A C T C G C C T A T T G T T A A 5′

The sequence has 2-fold symmetry regions; the two longest are separated by one turn of the DNA double helix.

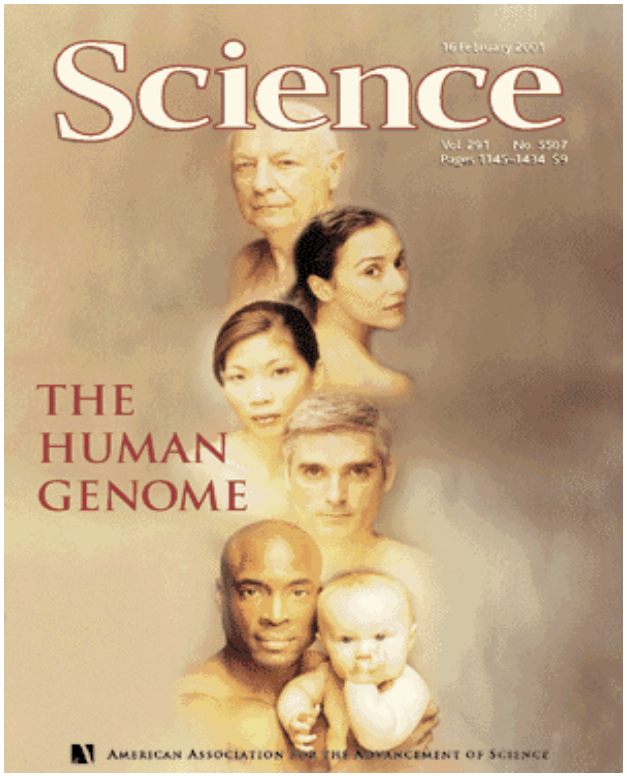
The lactose repressor selects one out of six million nucleotide sequences in the *Escherichia coli* genome and binds to it to prevent the expression of the genes for lactose metabolism. [illegible] bind again to the repressor, and is about 27 base-pairs long. Here we shall describe its sequence.

METHODS

Sonicated DNA Fragments. Sonicated [^{32}P]DNA fragments were made by growing a temperature-inducible lysogen of λ*c*1857p*lac*5S7 at 34° in a glucose–50 mM Tris·HCl or TES (pH 7.4) medium in 3 mM phosphate, heating at 42° for 15 min at a cell density of 4×10^8/ml, then washing and resuspending the cells at a density of 8×10^8/ml in the same medium with 0.1 mM phosphate. 100 mCi of neutralized $H_2{}^{32}PO_4$ was added to 10 ml of cells, and the incorporation was continued for 2 hr at 34°. The cells were washed, suspended in 2

Publication of Draft Version

International Human Genome Consortium

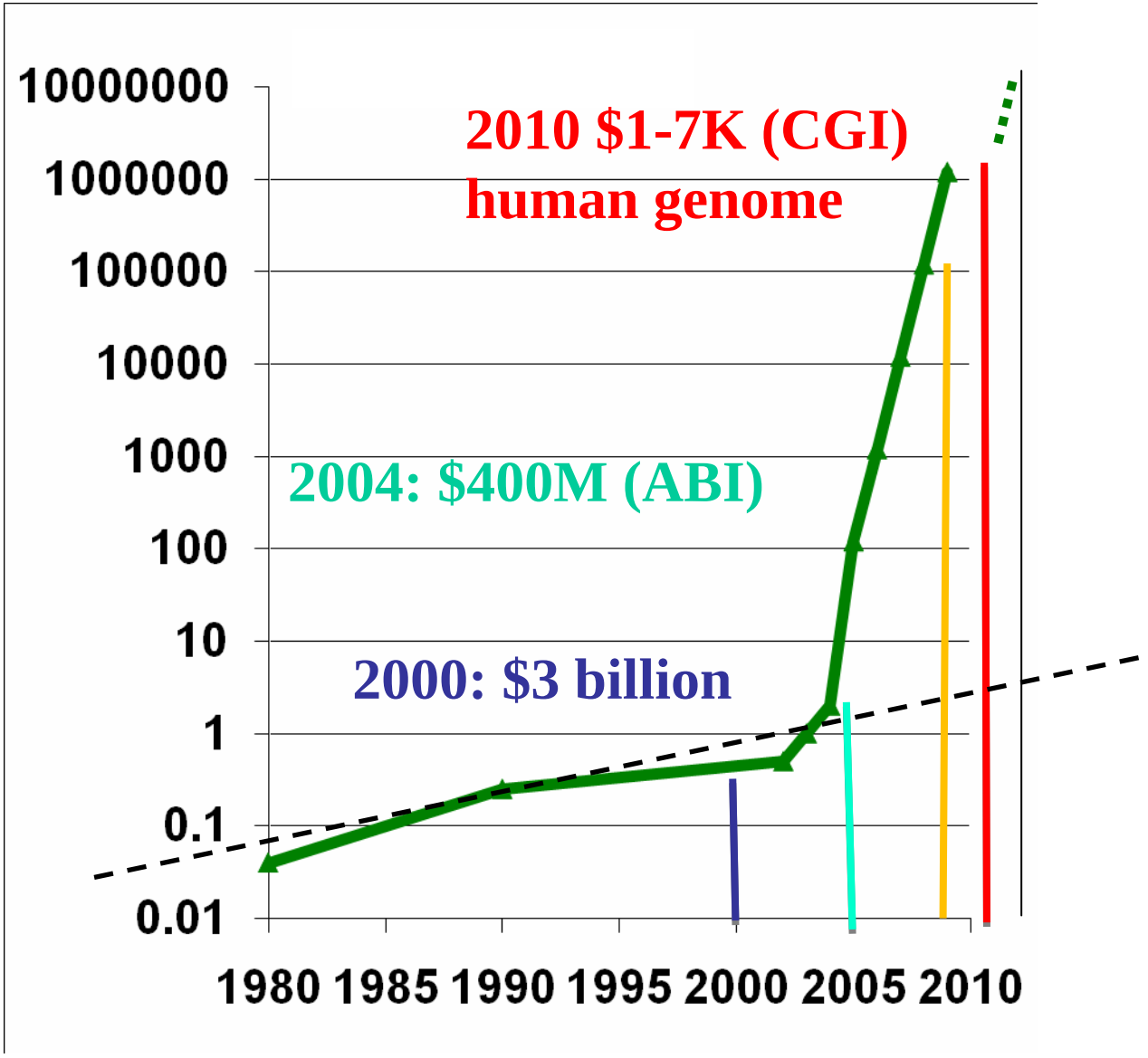


Celera

Genome$

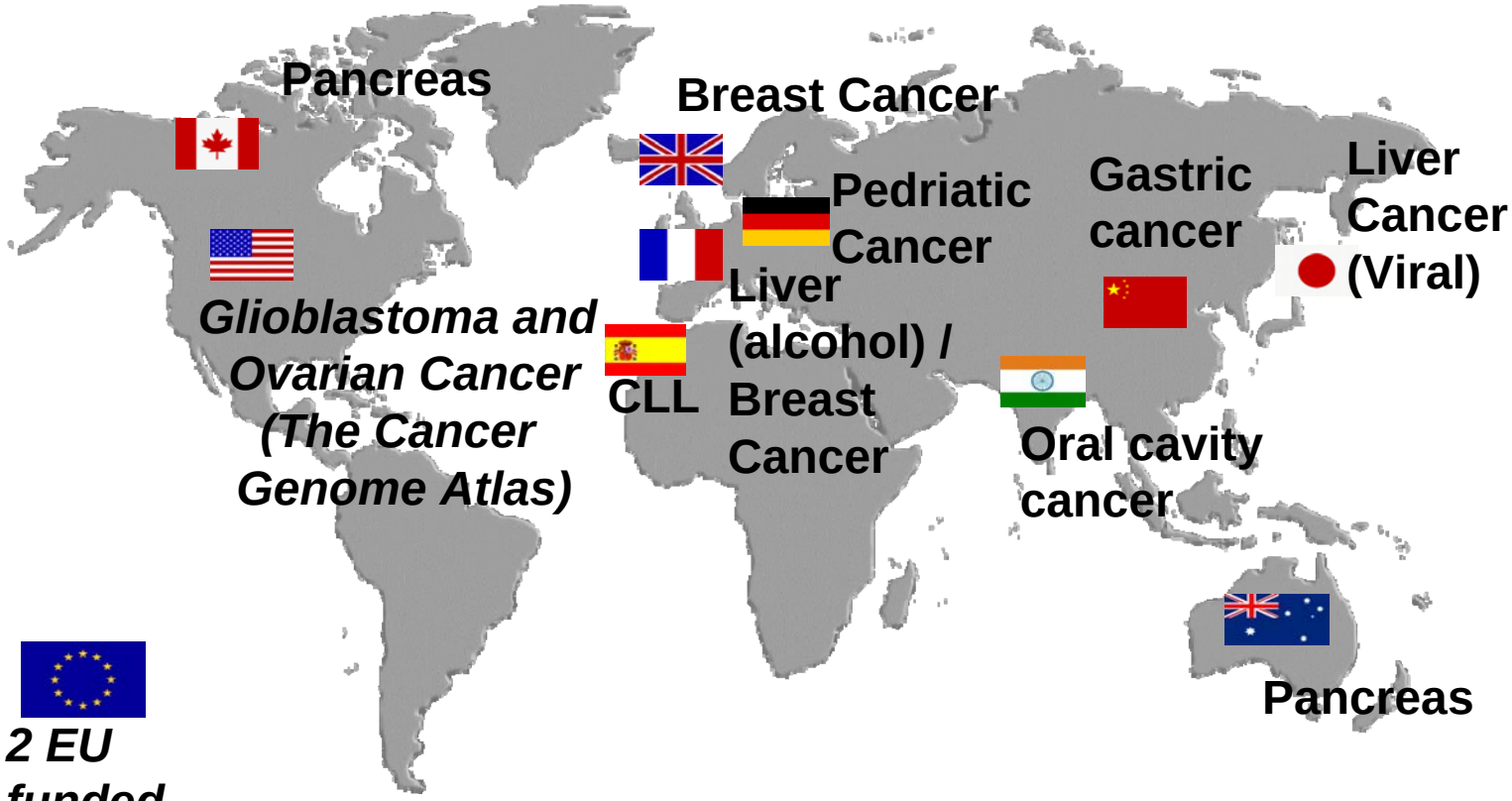
Factors of 10 since 2005

Moore's law 1.5x/yr for electronics



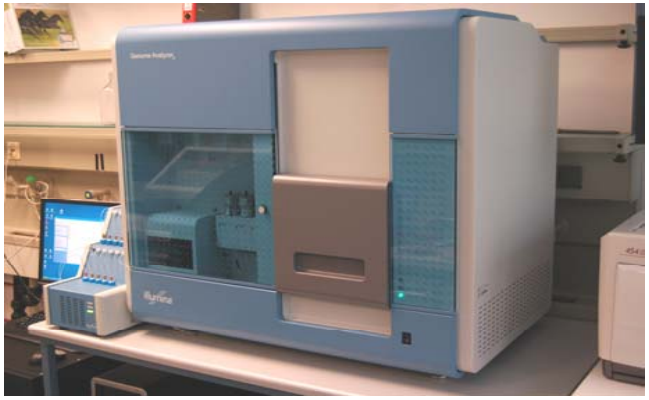
The 1000 genomes project: a catalogue of human polymorphism created using next generation sequencing

International Cancer Genome Consortium (ICGC)

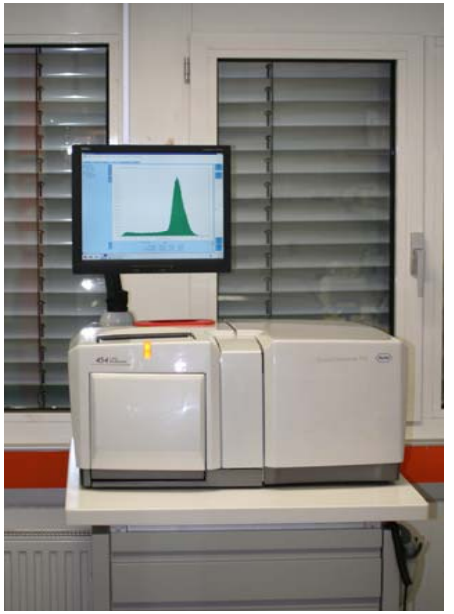


2 EU funded consortia

Sequencing Capacities at the MPI-MG



5 x Illumina GAII



3 x 454/FLX

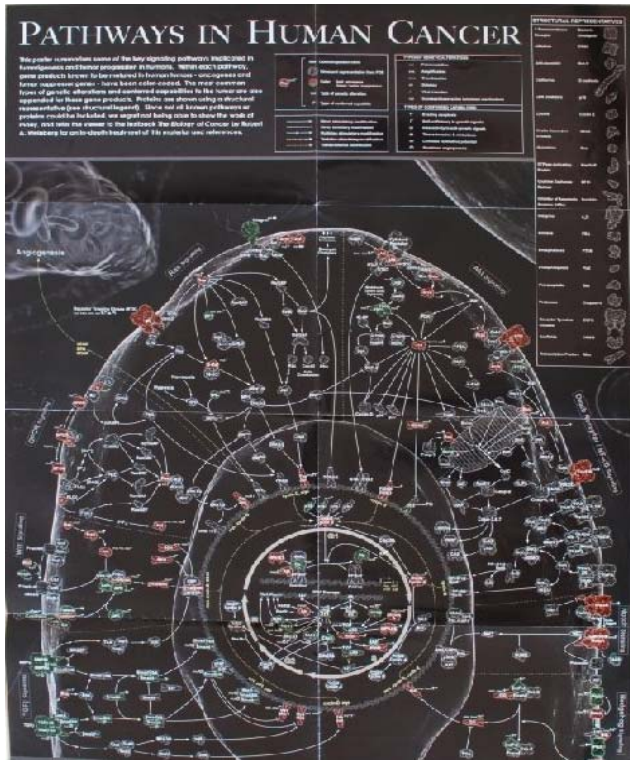
5 x SOLiD

1 x Polonator G007

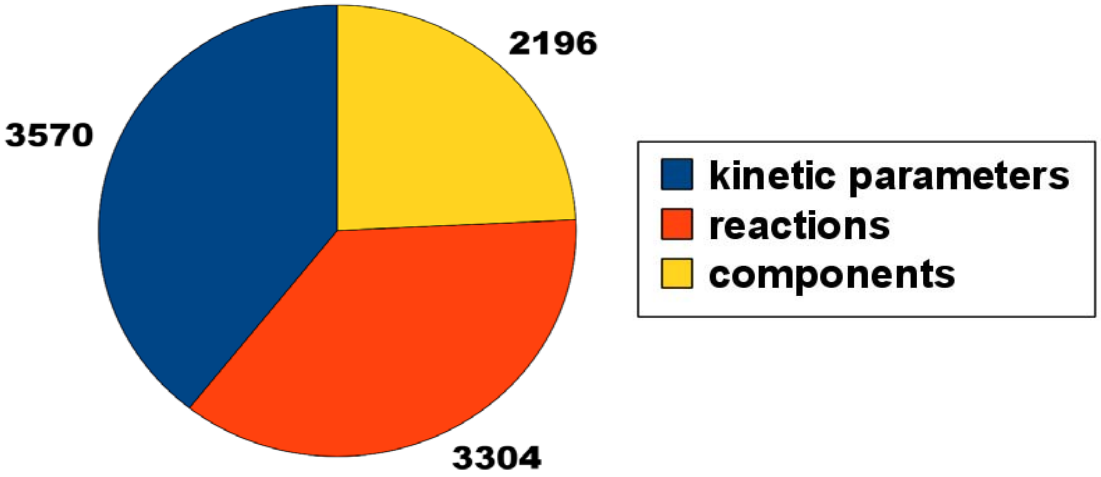
The MPI-MG is the largest second generation sequencing centre on the continent, technology, which will be directly transferred to Alacris Theranostics

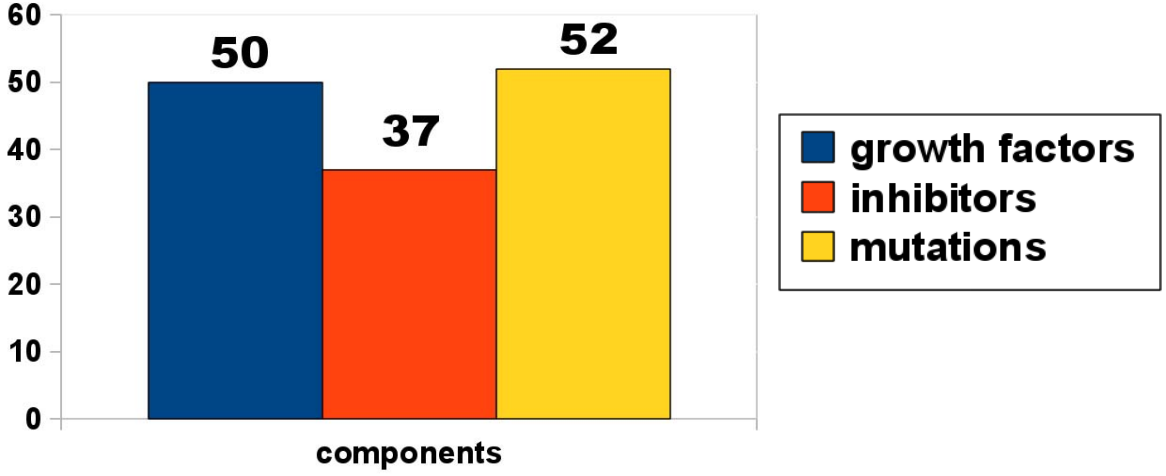
Christoph Wierling

Hanahan / Weinberg Cancer Model



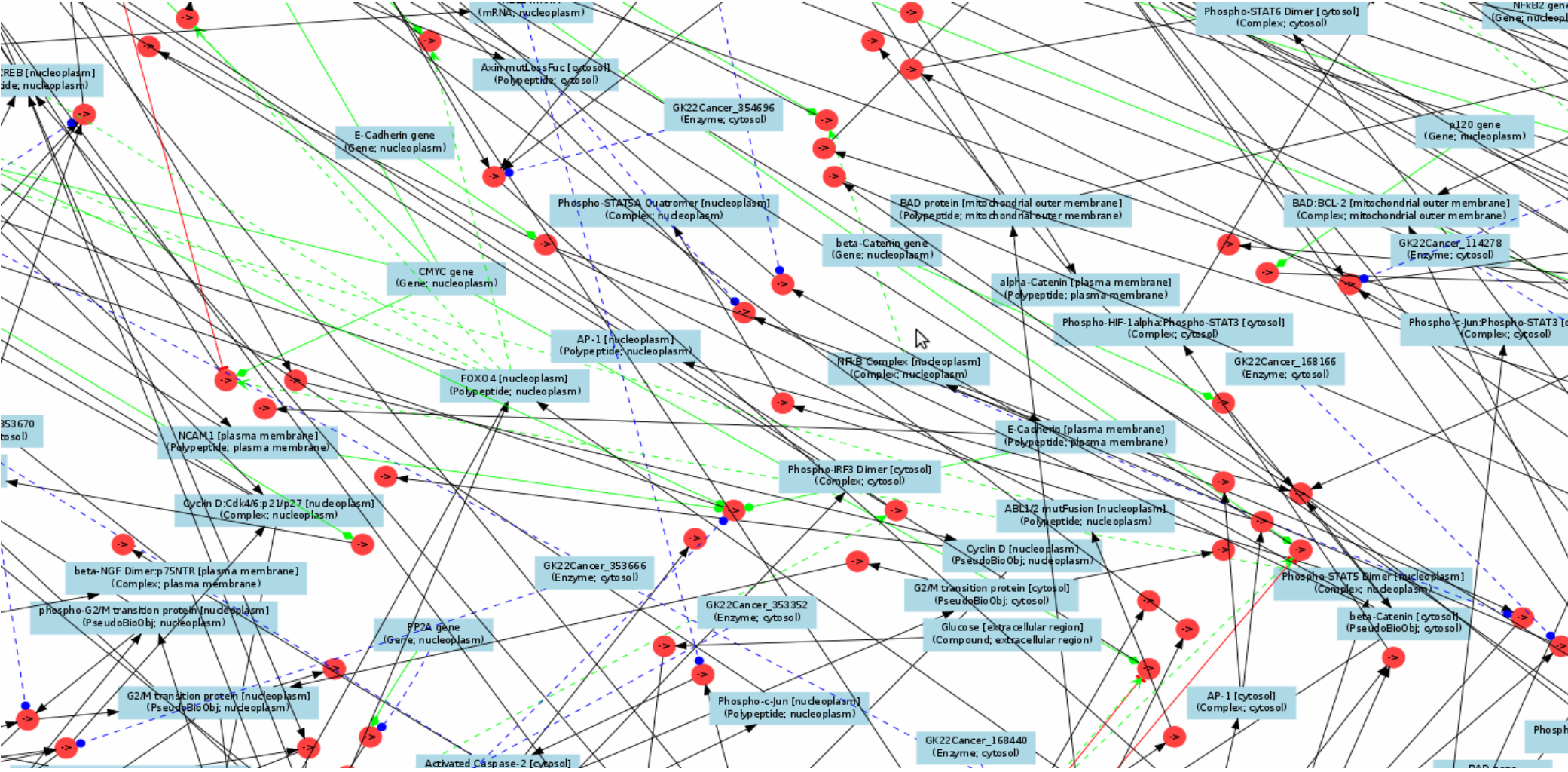
Weinberg, The Biology of Cancer (2007) Garland Science

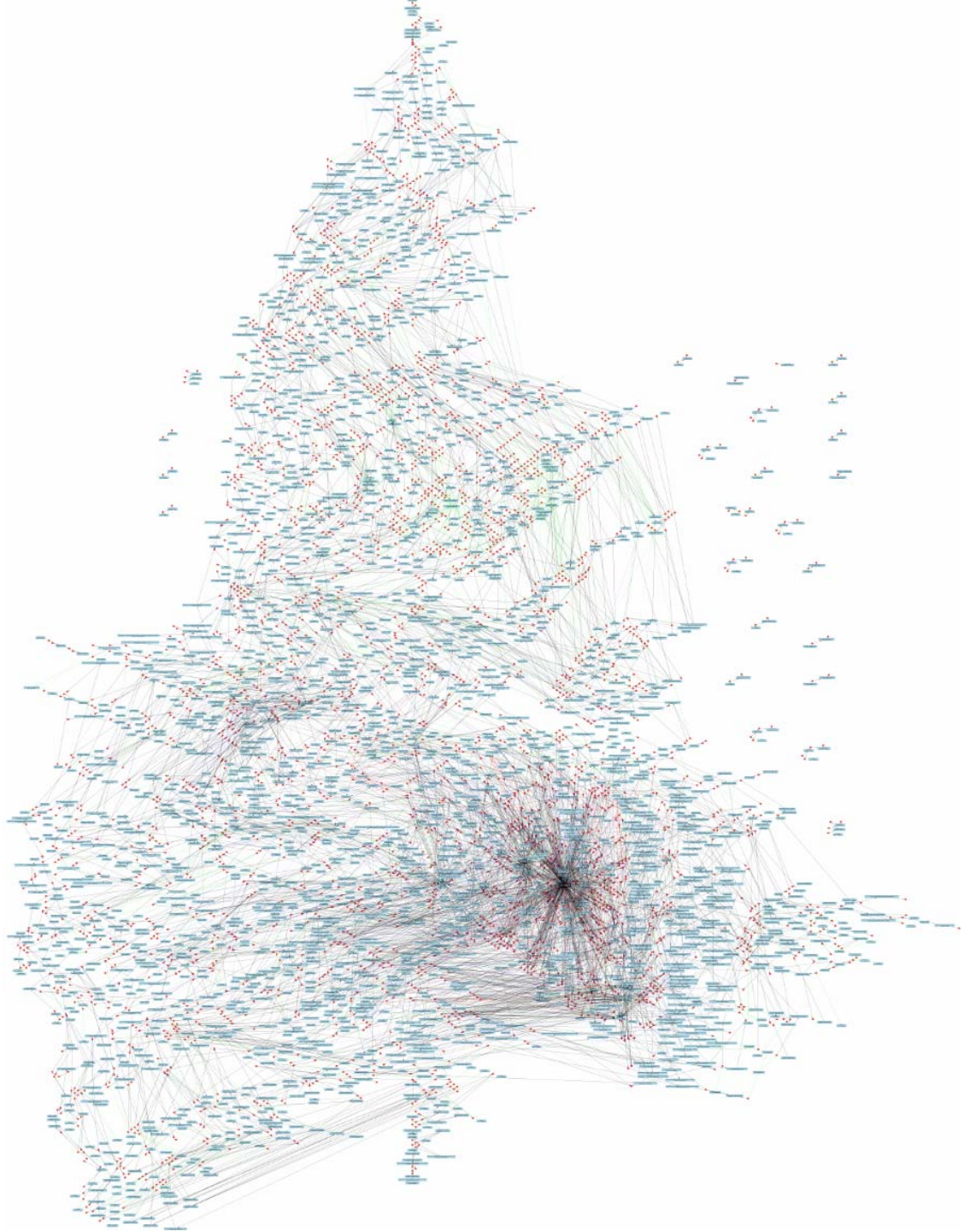




Pathways of the model

IGF-1 signaling	E-cadherin pathway
Cytokine signaling	Wnt signaling
BMP signaling	PLC signaling
TGFbeta signaling	EGF signaling
Hedgehog signaling	TLR3/TLR10 signaling
Notch signaling	GPCR pathway
TNF-alpha signaling	NGFR signaling
Fas signaling	Rb/E2F pathway
TRAIL signaling	DNA repair


(mRNA; nucleoplasm)
Axin mutLossFuc [cytosol]
(Polypeptide; cytosol)
GK22Cancer_354696
(Enzyme; cytosol)
E-Cadherin gene
(Gene; nucleoplasm)
Phospho-STAT5A Quatromer [nucleoplasm]
(Complex; nucleoplasm)
CMYC gene
(Gene; nucleoplasm)
AP-1 [nucleoplasm]
(Polypeptide; nucleoplasm)
FOXO4 [nucleoplasm]
(Polypeptide; nucleoplasm)
NCAM1 [plasma membrane]
(Polypeptide; plasma membrane)
Cyclin D:Cdk4/6:p21/p27 [nucleoplasm]
(Complex; nucleoplasm)
beta-NGF Dimer:p75NTR [plasma membrane]
(Complex; plasma membrane)
phospho-G2/M transition protein [nucleoplasm]
(PseudoBioObj; nucleoplasm)
G2/M transition protein [nucleoplasm]
(PseudoBioObj; nucleoplasm)
PP2A gene
(Gene; nucleoplasm)
GK22Cancer_353666
(Enzyme; cytosol)
GK22Cancer_353352
(Enzyme; cytosol)
Phospho-c-Jun [nucleoplasm]
(Polypeptide; nucleoplasm)
Activated Caspase-2 [cytosol]
Phospho-STAT6 Dimer [cytosol]
(Complex; cytosol)
p120 gene
(Gene; nucleoplasm)
BAD protein [mitochondrial outer membrane]
(Polypeptide; mitochondrial outer membrane)
BAD:BCL-2 [mitochondrial outer membrane]
(Complex; mitochondrial outer membrane)
beta-Catenin gene
(Gene; nucleoplasm)
GK22Cancer_114278
(Enzyme; cytosol)
alpha-Catenin [plasma membrane]
(Polypeptide; plasma membrane)
Phospho-HIF-1alpha:Phospho-STAT3 [cytosol]
(Complex; cytosol)
Phospho-c-Jun:Phospho-STAT3
(Complex; cytosol)
NFkB Complex [nucleoplasm]
(Complex; nucleoplasm)
GK22Cancer_168166
(Enzyme; cytosol)
E-Cadherin [plasma membrane]
(Polypeptide; plasma membrane)
Phospho-IRF3 Dimer [cytosol]
(Complex; cytosol)
ABL1/2 mutFusion [nucleoplasm]
(Polypeptide; nucleoplasm)
Cyclin D [nucleoplasm]
(PseudoBioObj; nucleoplasm)
G2/M transition protein [cytosol]
(PseudoBioObj; cytosol)
Glucose [extracellular region]
(Compound; extracellular region)
Phospho-STAT5 Dimer [nucleoplasm]
(Complex; nucleoplasm)
beta-Catenin [cytosol]
(PseudoBioObj; cytosol)
AP-1 [cytosol]
(Complex; cytosol)
GK22Cancer_168440
(Enzyme; cytosol)



The TREAT 1000 project

Making tomorrow's treatment available today

WELCOME TO THE TREAT 1000 WEBSITE!

TREAT1000 is an innovative project with the aim of bringing the benefits of genomic medicine to the cancer care of 1000 patients now. TREAT1000 was founded by top researchers and doctors from the Max Planck Institute of Molecular Genetics in Berlin, Harvard Medical School, the Charité Universitätsmedizin Berlin, Alacris Pharmaceuticals GmbH and CollabRx Inc. The project aim is to use a hybrid combination of funding sources, including patient and donor funding to fund applied research in patient treatment, research which will lead both to medical advances and to direct benefit for the patients involved in the project.

Every patient and every tumor is unique. Sequencing each patient's genome and their tumor genome will help their oncologists understand the specific mechanisms of tumor resistance and susceptibility for each patient's specific disease.

All data and conclusions generated in the project will be made publicly available through collaborations with the Personal Genome Project and with Health Commons.

More information about TREAT1000:

TREAT1000 workshop at Harvard Medical School, January 21st-22nd

Proof of Concept

Dr. Schlag with the first patient's tumor

Patient 1 Genome Sequencing

Sample	Type	Amount sequenced
Blood	Whole Genome	90 Gb
Tumor	Whole Genome	120 Gb
Blood	Exome	997 Mb
Tumor	Exome	892 Mb

Ca. 30x genome coverage for high-confidence scoring of mutations

- Heterogenous tumor with > 40,000 somatic mutations, mostly passengers !

Ca. 1,000 mutations in expressed proteins, mostly changing one amino acid

- One known cancer activating mutation: BRAF(V600E) observed in 50% of melanomas but not sufficient to promote tumorigenesis

Another 5 mutations in expressed genes model components for instance in PDGFD, PRKCB, SCRF1.... but the mutations effects are unknown :
being modeled for three possibilities (neutral,activating or loss of function)

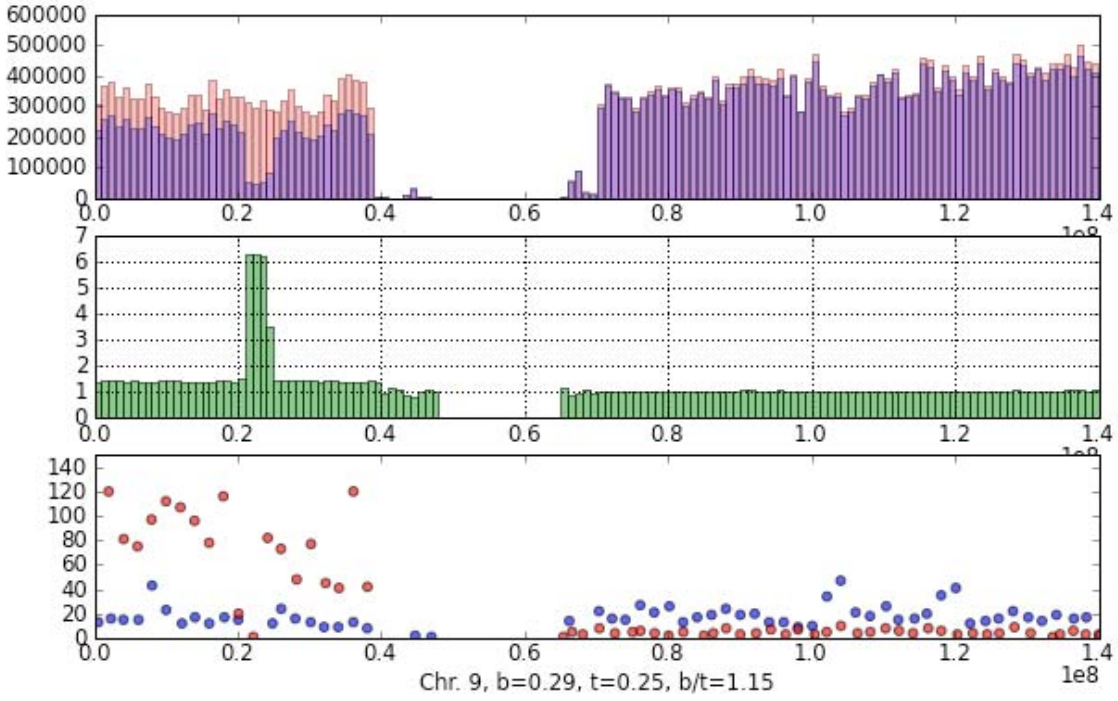
Identification of NOVEL driver mutations which will be implemented in the model)

Whole Genome Sequence provides virtual karyotype

(Example for tumor Chr.9) Identification of a large deletion of chr. 9

Gene Deletion

CDKN2A Cyclin-dependent kinase inhibitor 2A (ARF)
CDKN2B Cyclin-dependent kinase 4 inhibitor B



Amount of DNA per Mb shown for blood (red) and tumor (blue)

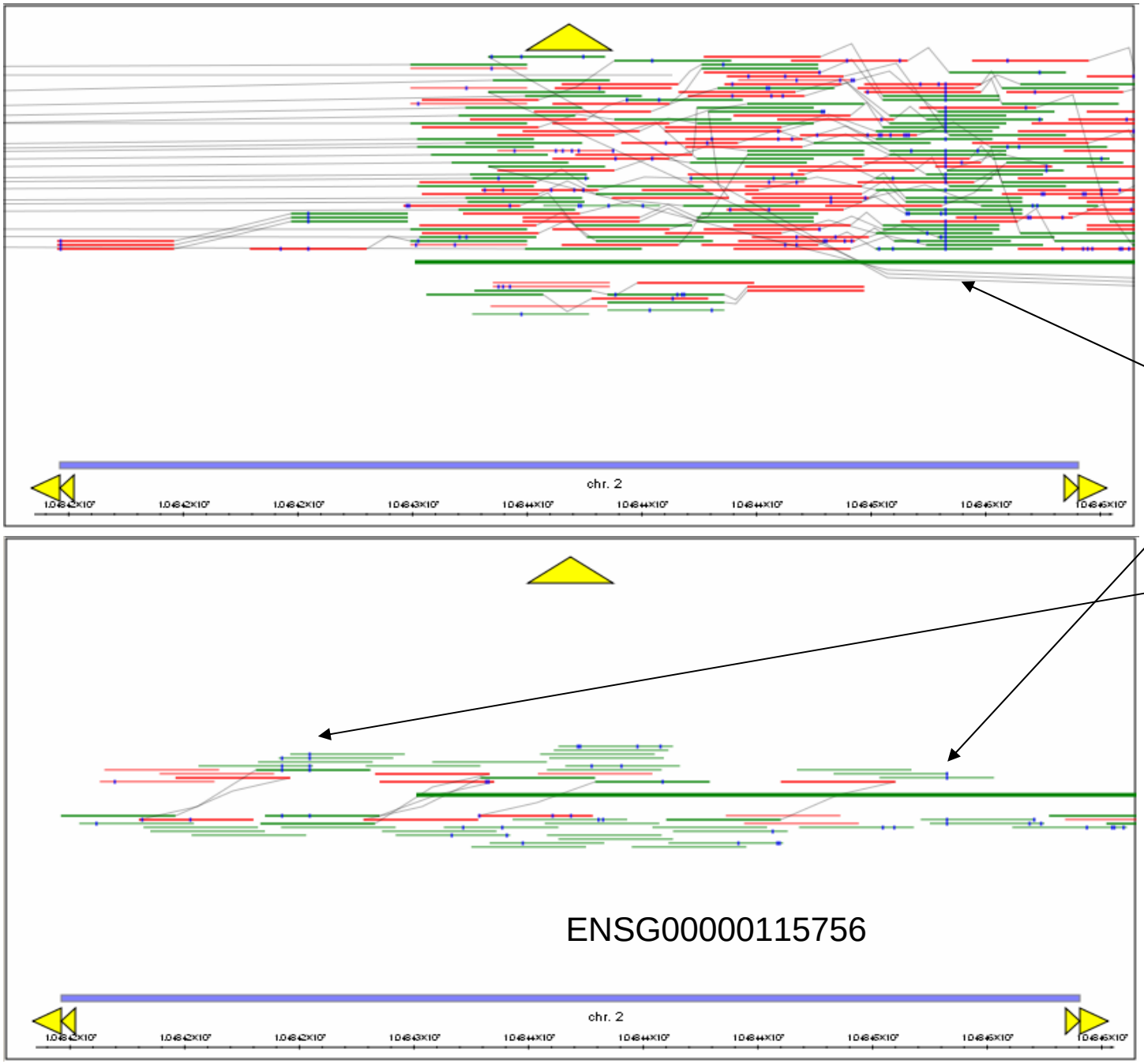
Ratio of reads (blood/tissue).

SNPs unique to blood (red), unique to tissue (blue). The higher number of blood-specific SNPs shows that this haplotype was lost in tumor

Patient 1 Transcriptome Sequencing

New protocol with RNA fragmentation		
Run/lane nb	Sample	Total.number_unique_hits
091211_EAS451_1	Tumor	28 859 047
091211_EAS451_2	Tumor	29 373 327
091211_EAS451_3	Tumor	29 456 853
091211_EAS451_4	Control	30 461 674
091211_EAS451_5	Control	29 622 024
Standard protocol with cDNA fragmentation		
Run/lane nb	Sample	Total.number_unique_hits
091113_EAS451_5	Tumor	21 995 533
091113_EAS451_6	Tumor	21 919 636
090522_EAS451_7	Tumor	11 277 659
090714_EAS451_1	Stem cells-CD133+	16 468 893
090714_EAS451_2	Stem cells-CD133-	21 808 807
091106_EAS451_6	Control	24 766 112

Tumor tissues and stem cells + control melanocyte sequenced

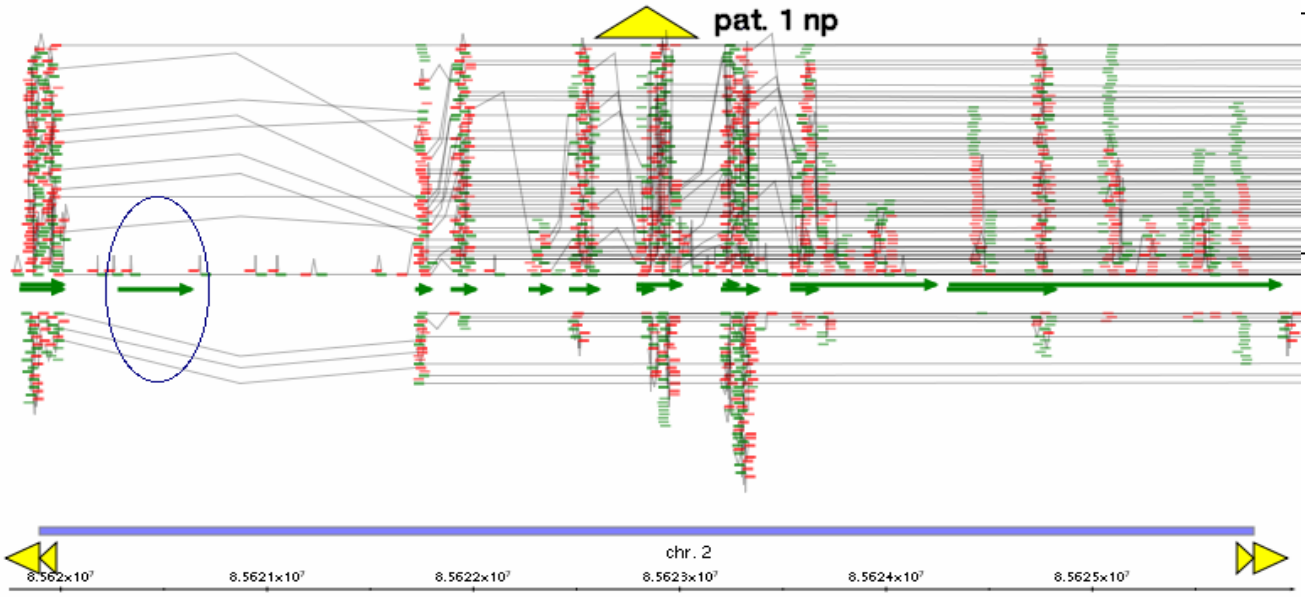


Cross-validated homozygous SNP

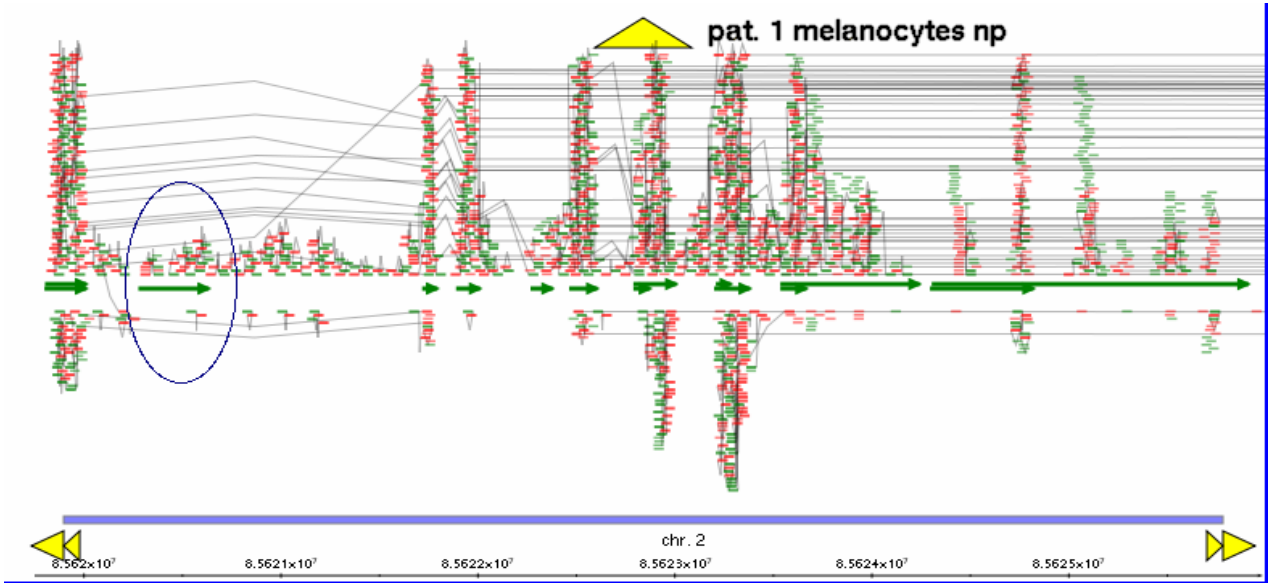
Heterozygous exome SNP (in splicing region) looks as homozygous in transcriptome

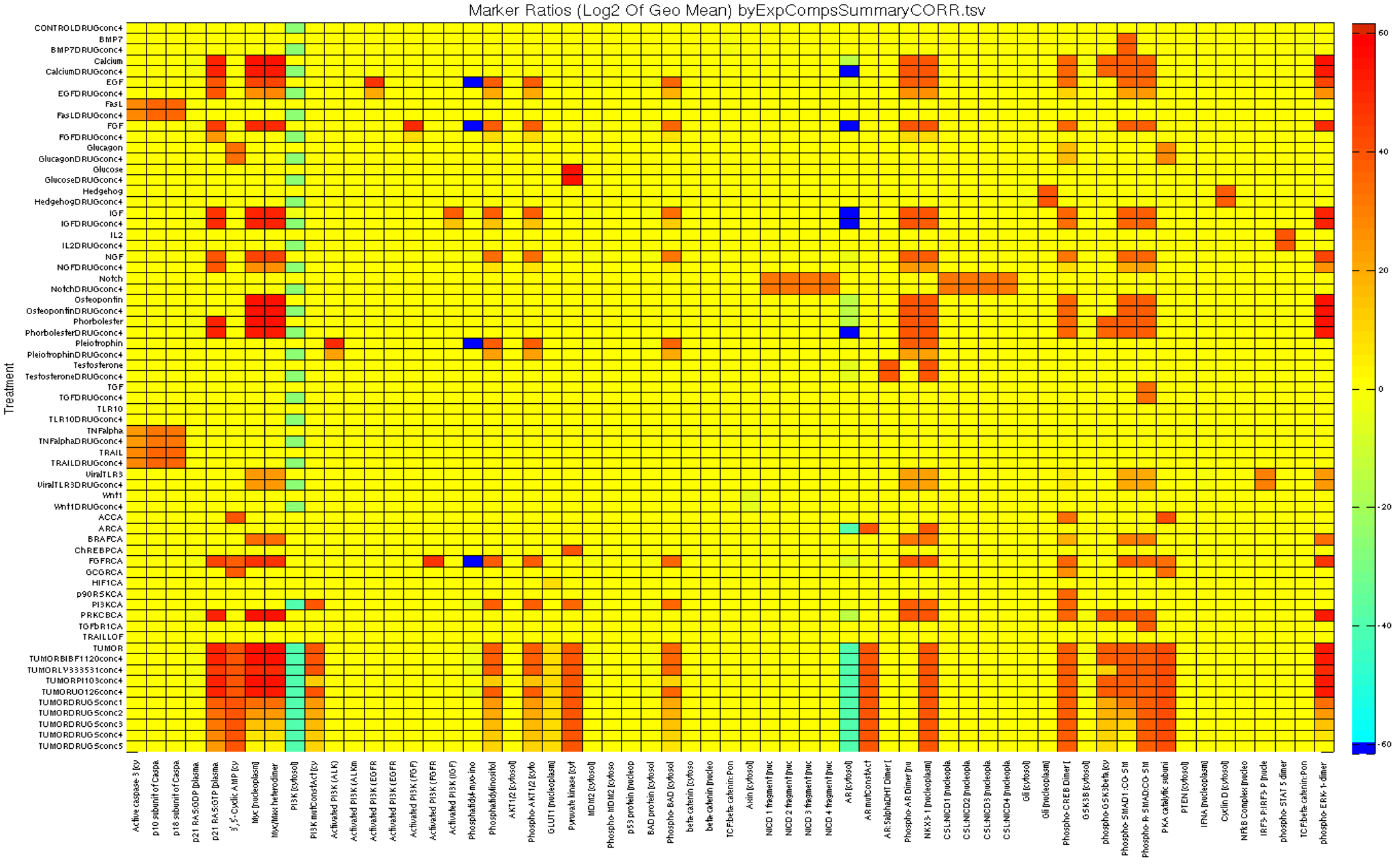
MAT2A

Pat1



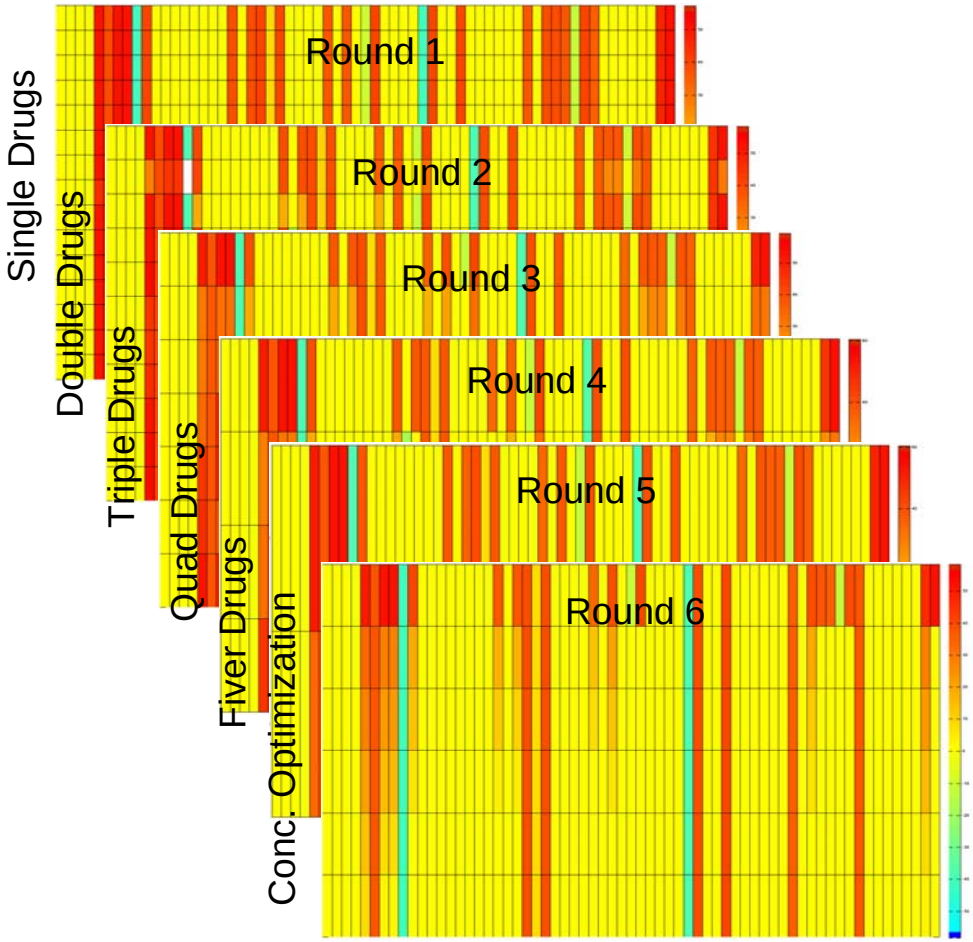
Melanocytes

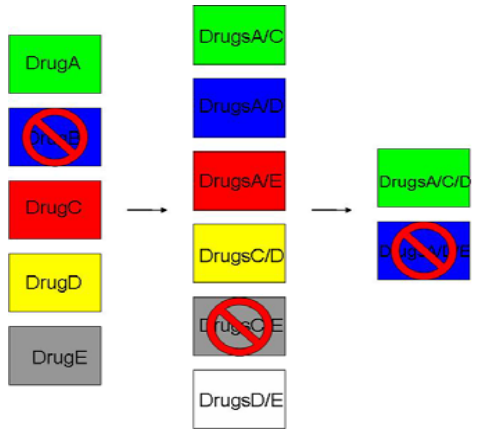


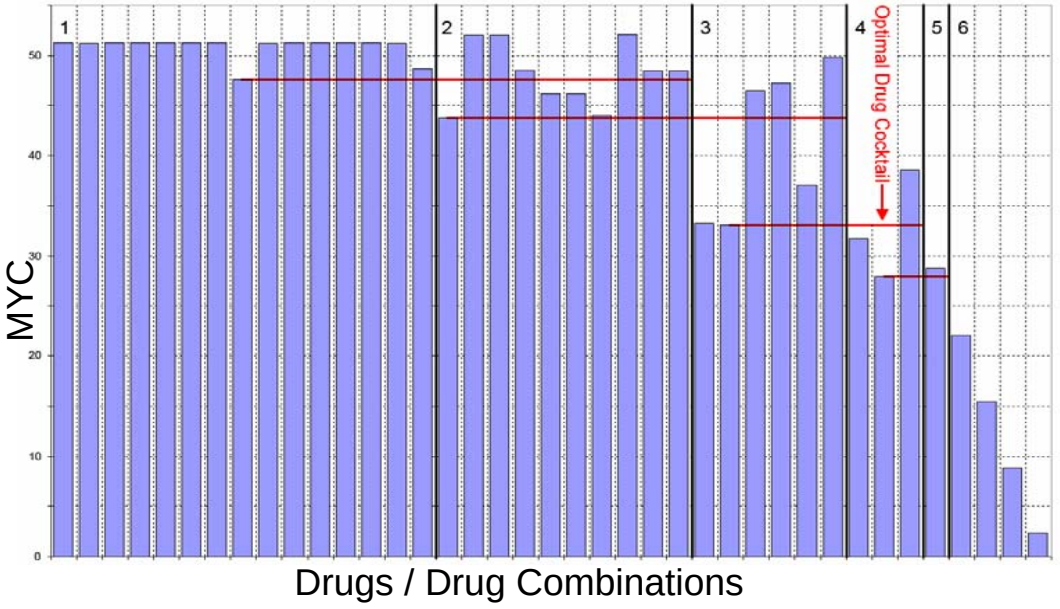


The model of patient 1

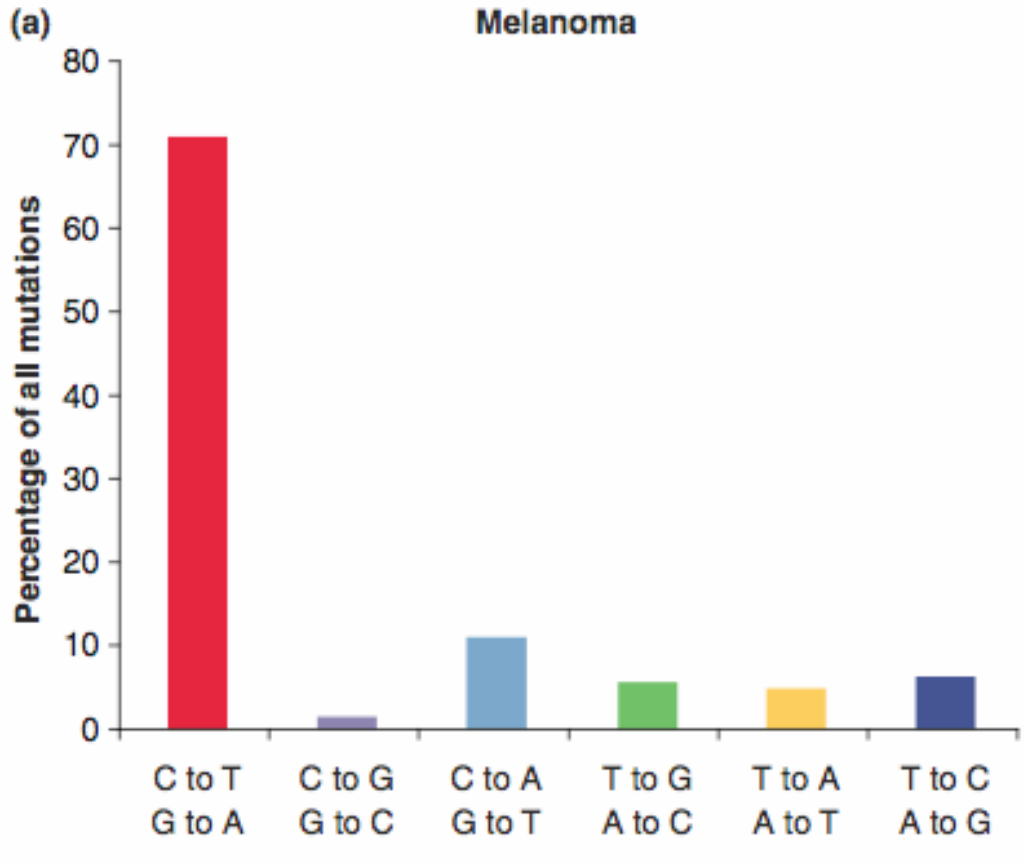
Finding a Patient 1 Specific Optimal Drug Treatment

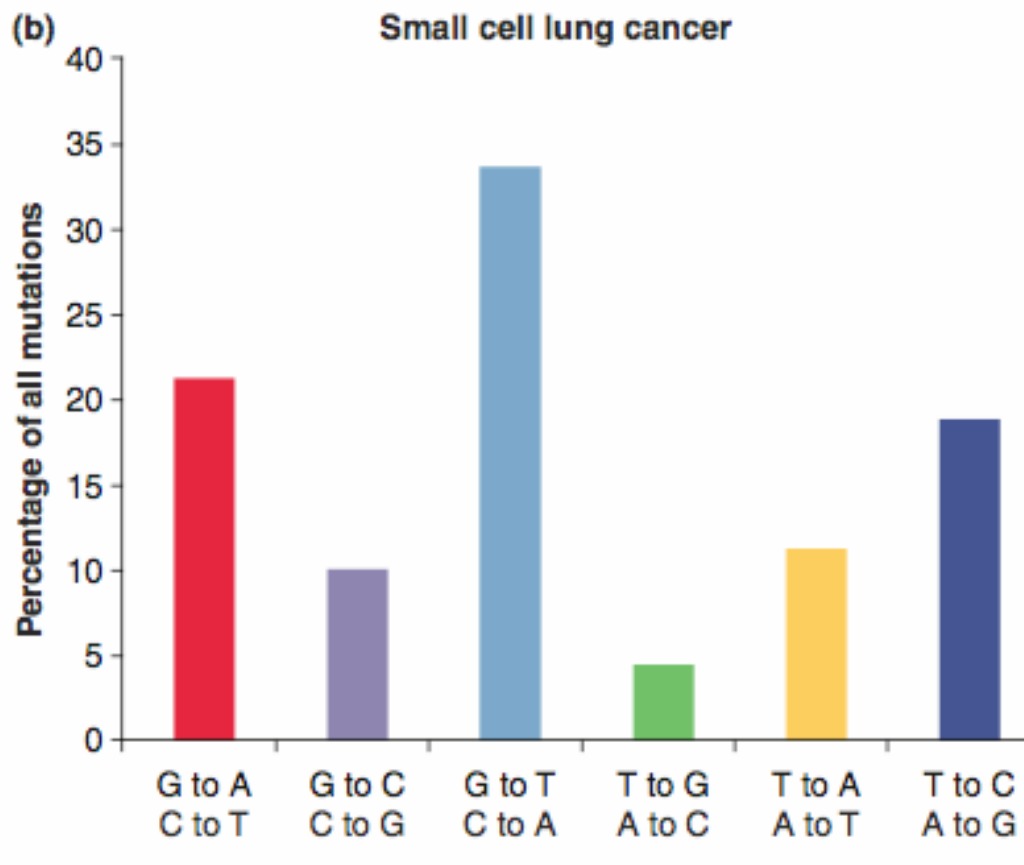






Stepwise optimisation of the therapy for an individual patient


(a)
Melanoma
Percentage of all mutations
80
70
60
50
40
30
20
10
0
C to T
G to A
C to G
G to C
C to A
G to T
T to G
A to C
T to A
A to T
T to C
A to G


(b)
Small cell lung cancer
Percentage of all mutations
40
35
30
25
20
15
10
5
0
G to A
C to T
G to C
C to G
G to T
C to A
T to G
A to C
T to A
A to T
T to C
A to G

Detection of environmental effects

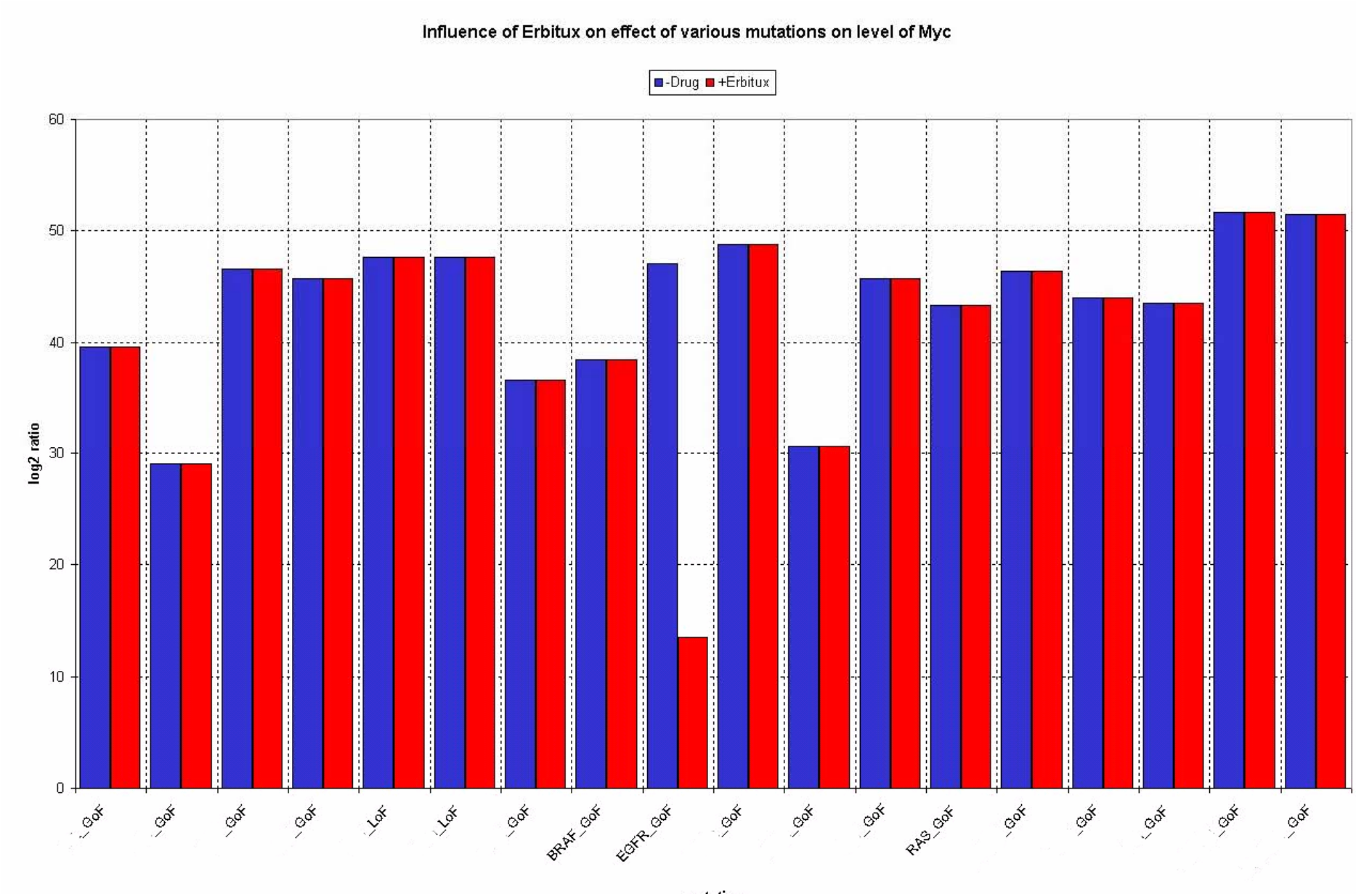
A) ‘mutagen fingerprints’

B) Statistically significant increase in specific molecular subtypes

C) Mechanistic predictions (e.g. tumor promoters)

D) In-silico tumor formation?

Modeling the effects of mutations, which affect response to EGFR inhibitors



- **PacBio has been quoted that, by 2013, their technology will be able to give a ‘raw’ human genome sequence in less than 3 min, and a complete high-quality sequence in just 15 min**

Modelling the response of cancer stem cells?

Cancer, a curable disease?

The ‚red Queen‘ approach to cancer treatment: Modelling cancer progression.

Cancer, a chronic disease‘

IT Future of Medicine

Setting the Scene

Market

"NEW YORK, Dec. 8, 2010 /PRNewswire/ -- **Spending on healthcare** among the **OECD** (i) **countries** and **BRIC** nations of Brazil, Russia, India and China will grow by 51 percent **between 2010 and 2020**, amounting to a **cumulative total of more than $71 trillion**, according to estimates from PwC's Health Research Institute. Health spending in these areas is rising faster than gross domestic product, magnifying gaps in budget deficits and spurring governments to look to the private sector for ways to get a better value for taxpayers' money."

"**Worldwide IT Spending** outperformed expectations in 2010, reaching **$1.5 Trillion**", according to IDC's Worldwide Black Book from February 09, 2011 08:03 AM Eastern Time

12 Million new cancer cases every year ~1 genome/second?

Computing requirements:

Sequence analysis: 100 cores ~1 day/genome: 1e7 core seconds

Modelling: 1000 conditions, 1000 positions in the body, 1000 MC runs, 1000 core seconds/run: 10e12 core seconds

1 Blue Gene L: ~10e5 cores (up to 65,536 nodes with 2 Power PC processors)

~1 Blue Gene L equivalent per patient?

Storage:

~1 Terabyte per genome

30-100 Exabytes for data storage for all cancer patients?

Multi-Zetabytes to Yottabytes for ITFoM?

As of May 2009, the size of the world's total digital content has been roughly estimated to be 500 billion Gigabytes (500 Exabytes) (From the Wiki Exabyte entry)

Metabolomics

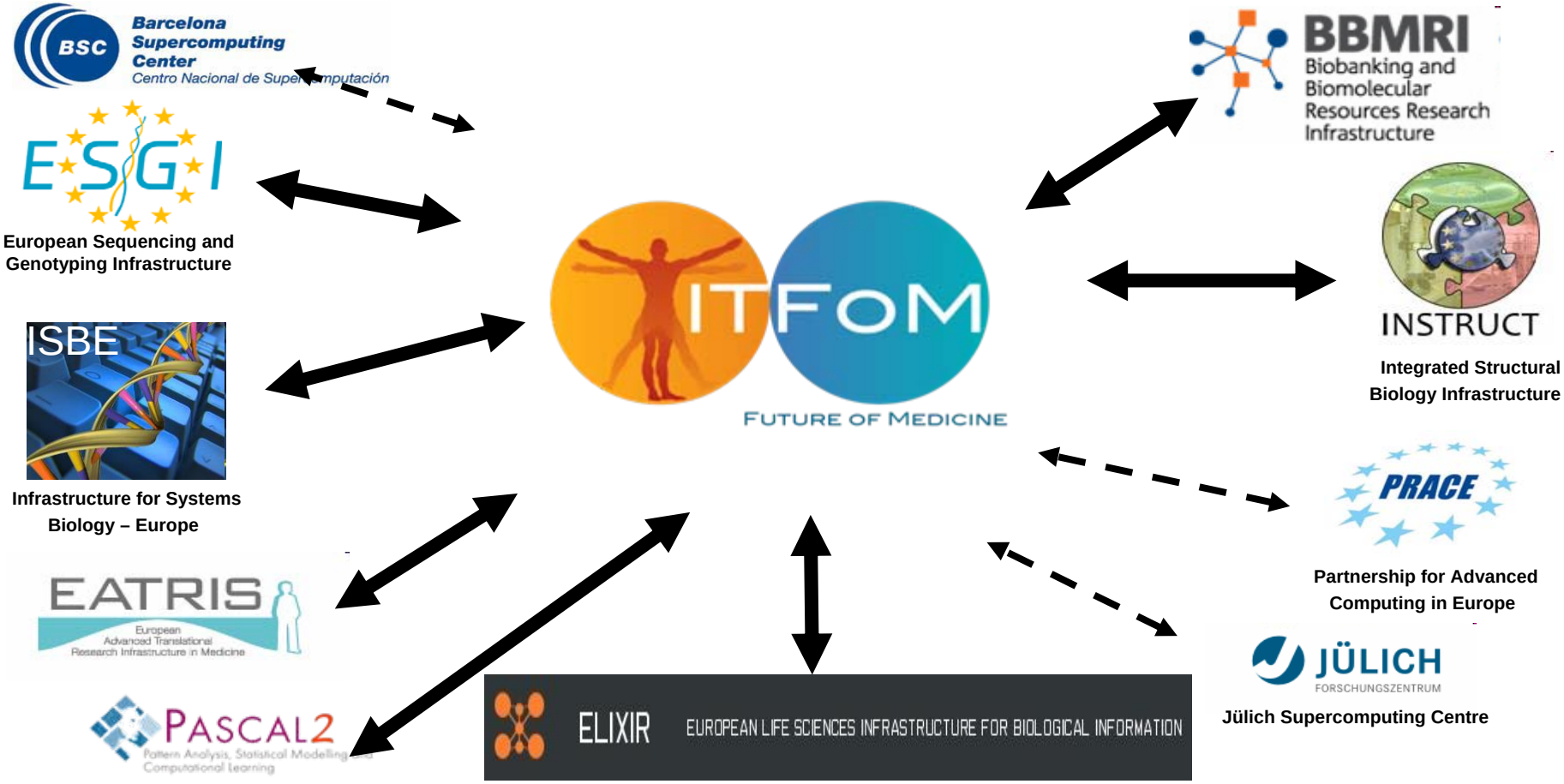
NMR-based metabolomics

Metabolic fingerprint

24 Partners

- Max Plank Institut for Molecular Genetics
- Medical University Graz
- University College London
- Free University of Amsterdam
- University of Manchester
- Maastricht University
- EMBL
- Wellcome Trust Sanger Institute
- Kungliga Tekniska högskolan
- Imperial College London
- CIRMMP
- International Prevention Research Institute
- Uppsala University
- University of Luxembourg
- University of Leicester
- HARVARD Medical School
- University of Auckland
- Universite de Geneve
- Centro Nacional De Análisis Genómico
- Siemens
- Alacris Theranostics GmbH
- Charite Universitätsmedizin Berlin
- Illumina
- Commissariat a l'energie atomique et aux energies alternatives

Relations to Infrastructures



The project outcomes will enable the prediction of health, disease, therapy and its effects for individual patients and through application in the clinic will change the future of medicine.

For more information:
Website: **http://www.itfom.eu**
Email: **info@itfom.eu**
Twitter: **@itfom**
Facebook: **I.T. Future of Medicine**
LinkedIn**: IT Future of Medicine**