AI applications for analysis ofmulti ‘Omics’ data for identification of personalized driver pathways and Cancer therapy candidates

Uğur Sezerman

Acıbadem Üniversitesi

ACIBADEM
ÜNİVERSİTESİ

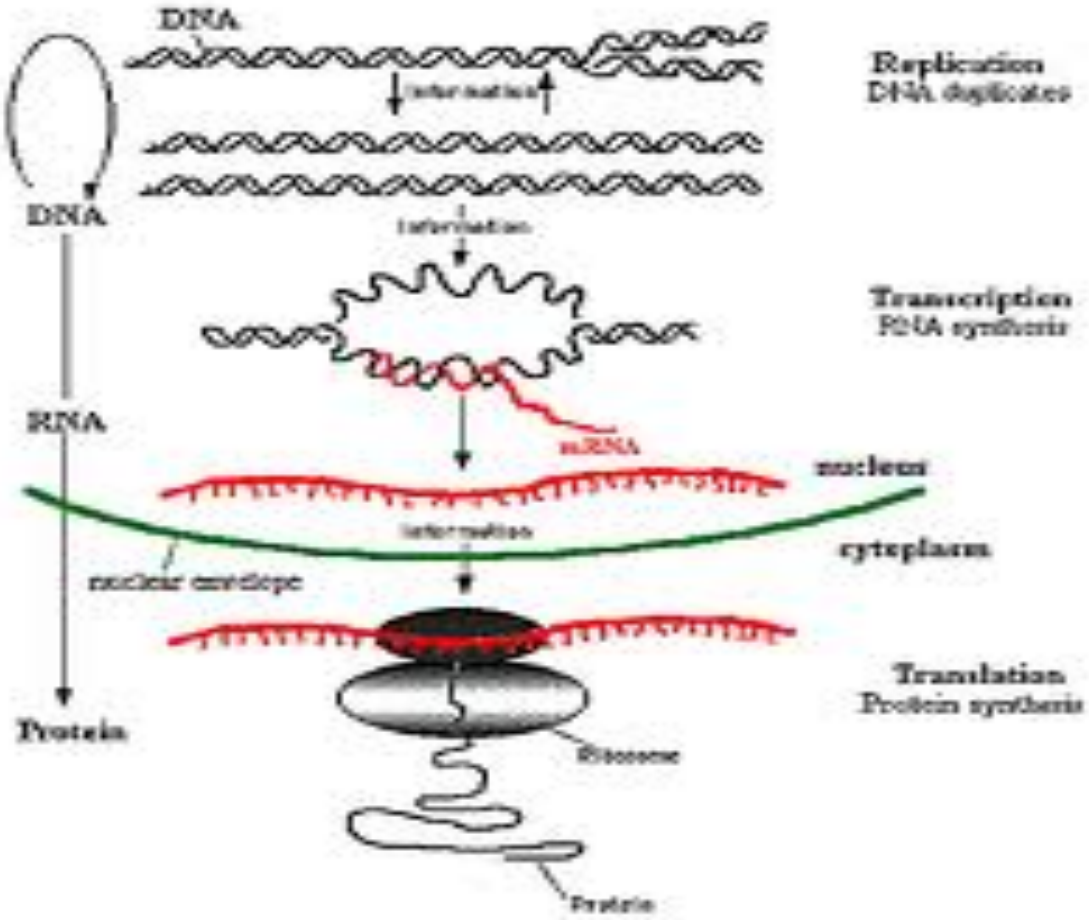
Goals:

- identify all the approximate 30,000 genes in human DNA,
- determine the sequences of the 3 billion chemical base pairs that make up human DNA,
- store this information in databases,
- improve tools for data analysis,
- transfer related technologies to the private sector, and
- address the ethical, legal, and social issues (ELSI) that may arise from the project.

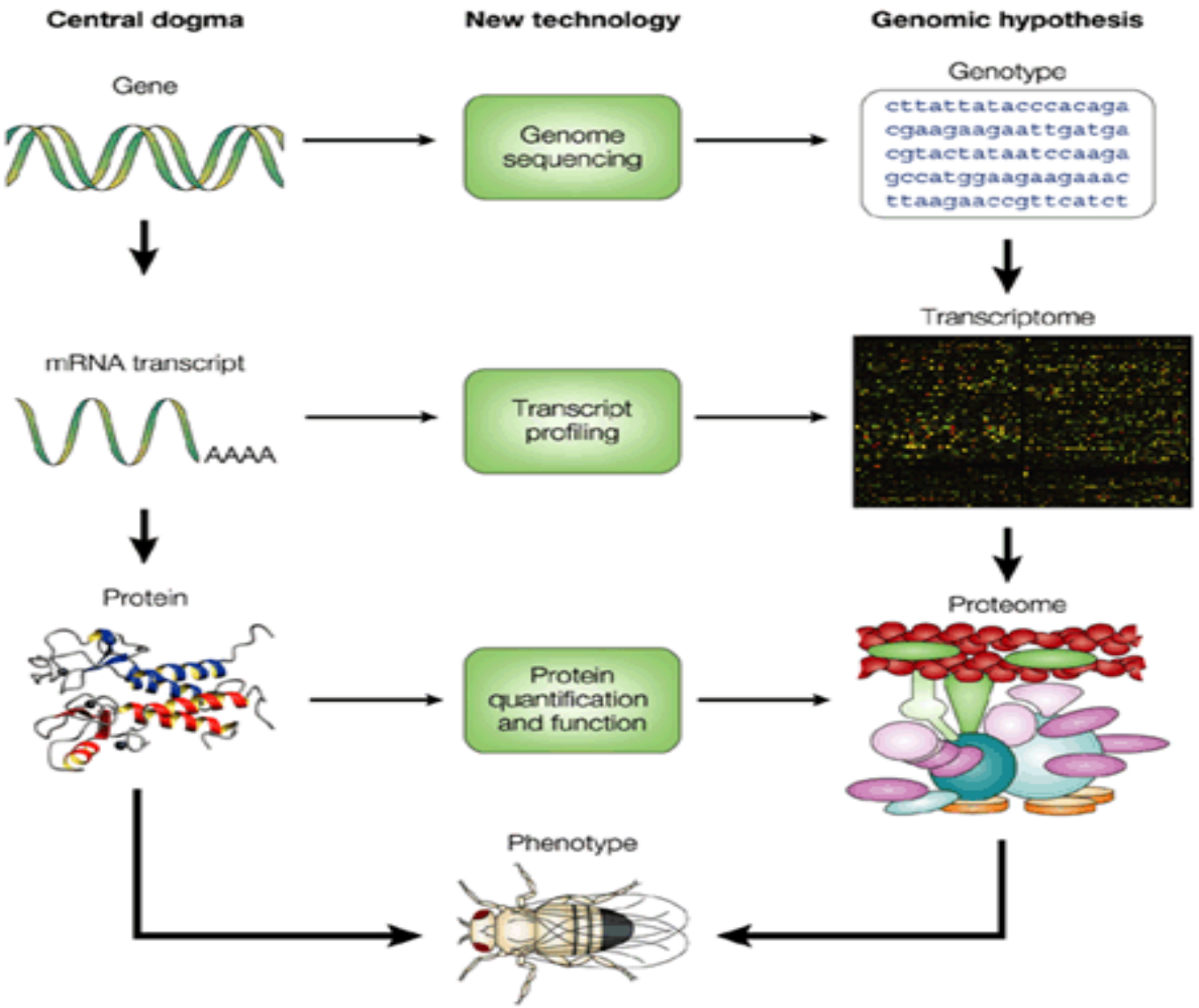
Milestones:

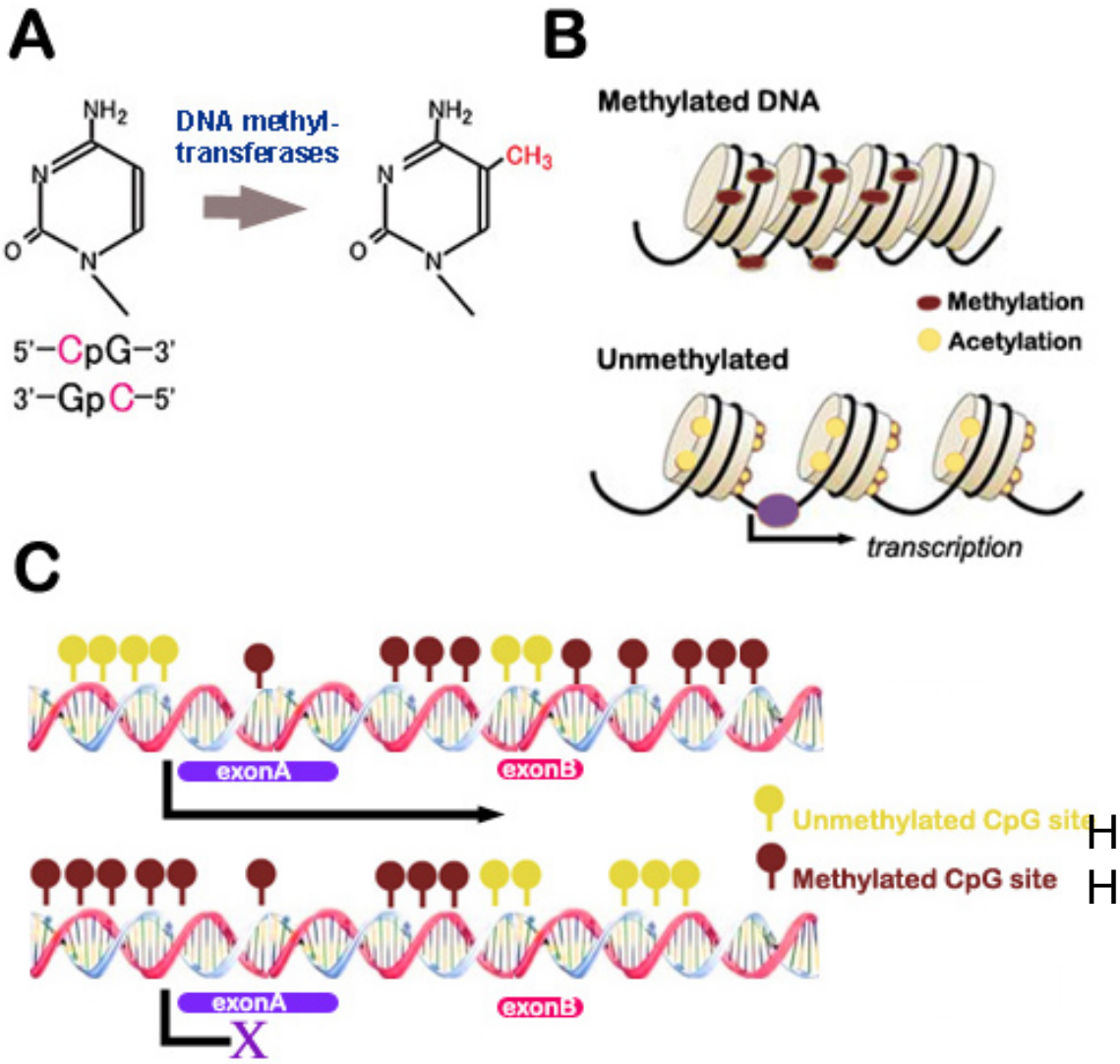
- 1990: Project initiated as joint effort of U.S. Department of Energy and the National Institutes of Health
- June 2000: Completion of a working draft of the entire human genome (covers >90% of the genome to a depth of 3-4x redundant sequence)
- February 2001: Analyses of the working draft are published
- April 2003: HGP sequencing is completed and Project is declared finished two years ahead of schedule

http://doegenomes.org
http://www.sanger.ac.uk/HGP/overview.shtml U.S. Department of Energy Genome Programs, Genomics and Its Impact on Science and Society, 2003



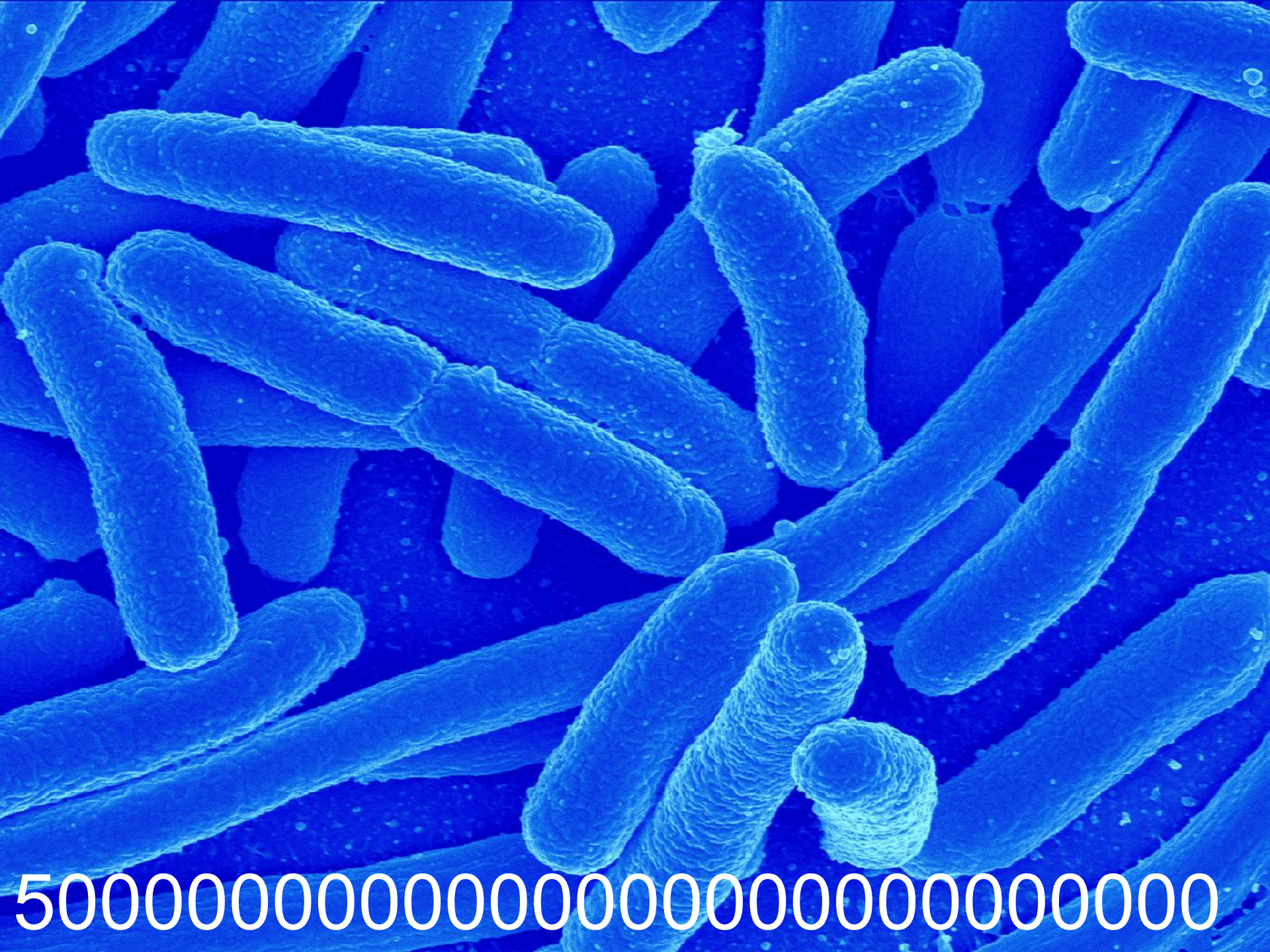
The Central Dogma of Molecular Biology

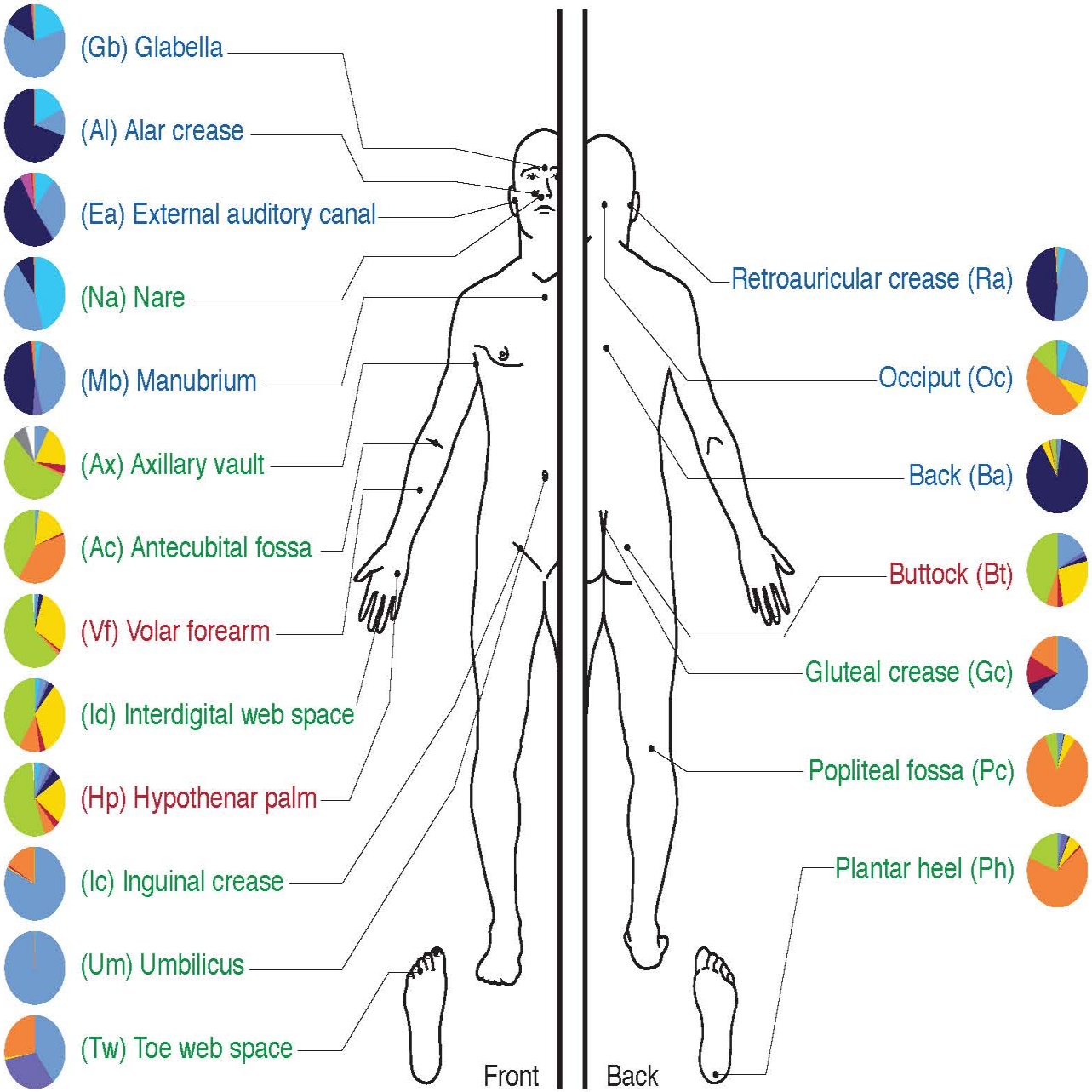

Central dogma
New technology
Genomic hypothesis
Gene
Genome sequencing
Genotype
cttattatacccacaga
cgaagaagaattgatga
cgtactataatccaaga
gccatggaagaagaaac
ttaagaaccgttcatct
mRNA transcript
AAAA
Transcript profiling
Transcriptome
Protein
Protein quantification and function
Proteome
Phenotype
Nature Reviews | Genetics

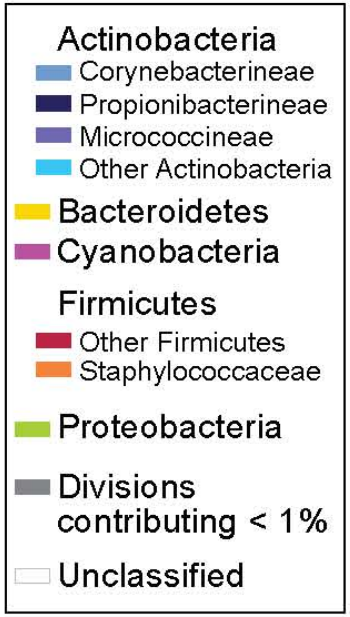


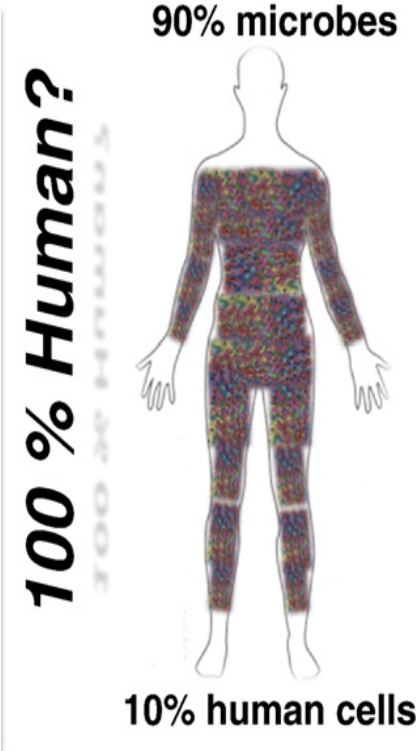
Hypomethylation
Hypermethylation

http://www.cellscience.com/reviews7/Taylor1.jpg


5000000000000000000000000000







We Are Really More Bug than

GUT MICROBIOTA

10^{13} -10^{14} microbes

1000- 35000 of species (most of them are still to be identified)

Weight – **3 to 5 lbs**

Genome – **150 fold of our Genome**

Bacteroides, Prevotella, Fusobacterium, Eubacterium, Ruminococcus, Peptococcus, Peptostreptococcus, Bifidobacterium. Escherichia* and *Lactobacillus.

Bacteroides alone constitute about 30% of all bacteria in the gut.....

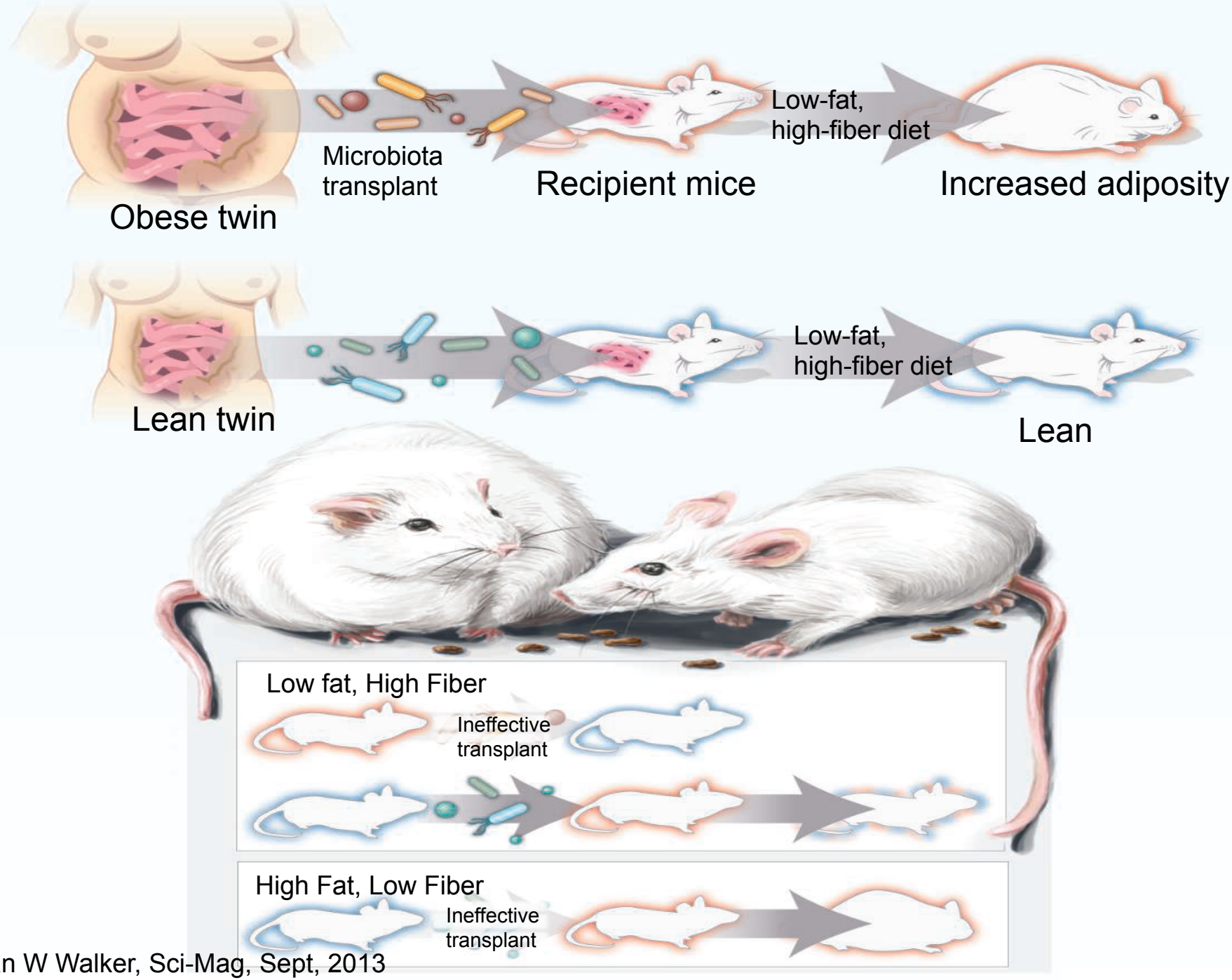
Carbohydrate fermentation and absorption
Digest starch, plant fiber, pectin into SCFAs (short chain fatty acids) viz. acetic acid, propionic acid, butyric acid. Digest proteins like collagen, elastin.

Repression of pathogenic microbial growth
Competition for nutrition, (ruminococus and prevettella) attachment. Produce bacteriocins , Lactic acid.Also Bacillus strains produces Bacilysin which kills closteridium botullinum

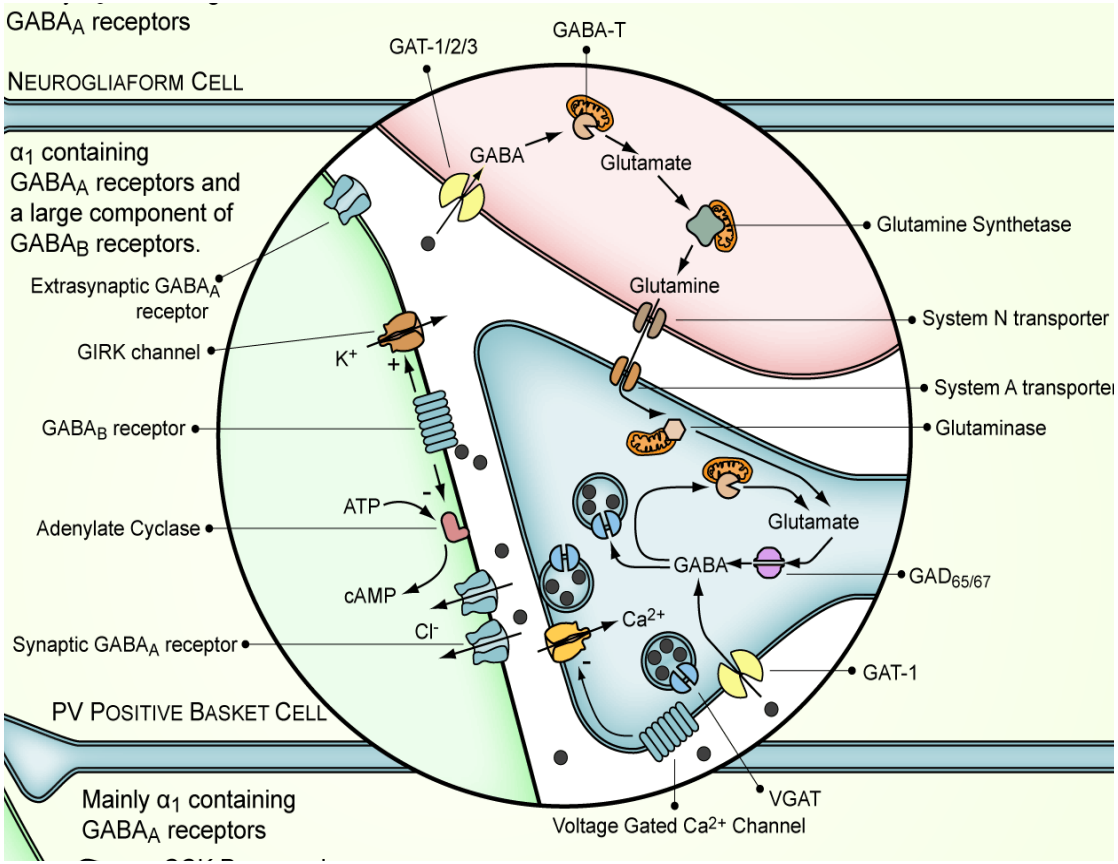
Metabolic function
HCA (heterocyclic amines)

Preventing inflammatory bowel disease
SCFAs prevent IBD

Preventing allergy
Allergies = *C. difficile* and *S. aureus* > *Bacteroides* and *Bifidobacteria*



Alan W Walker, Sci-Mag, Sept, 2013



Lactobacillus spp. and Bifidobacterium spp. produce GABA

GABA's natural function is to reduce the activity of the neurons to which it binds. GABA neutralizes the overexcited neurons. (anti-stress drug : Benzodiazepine)

AI/ML in Translational Medicine

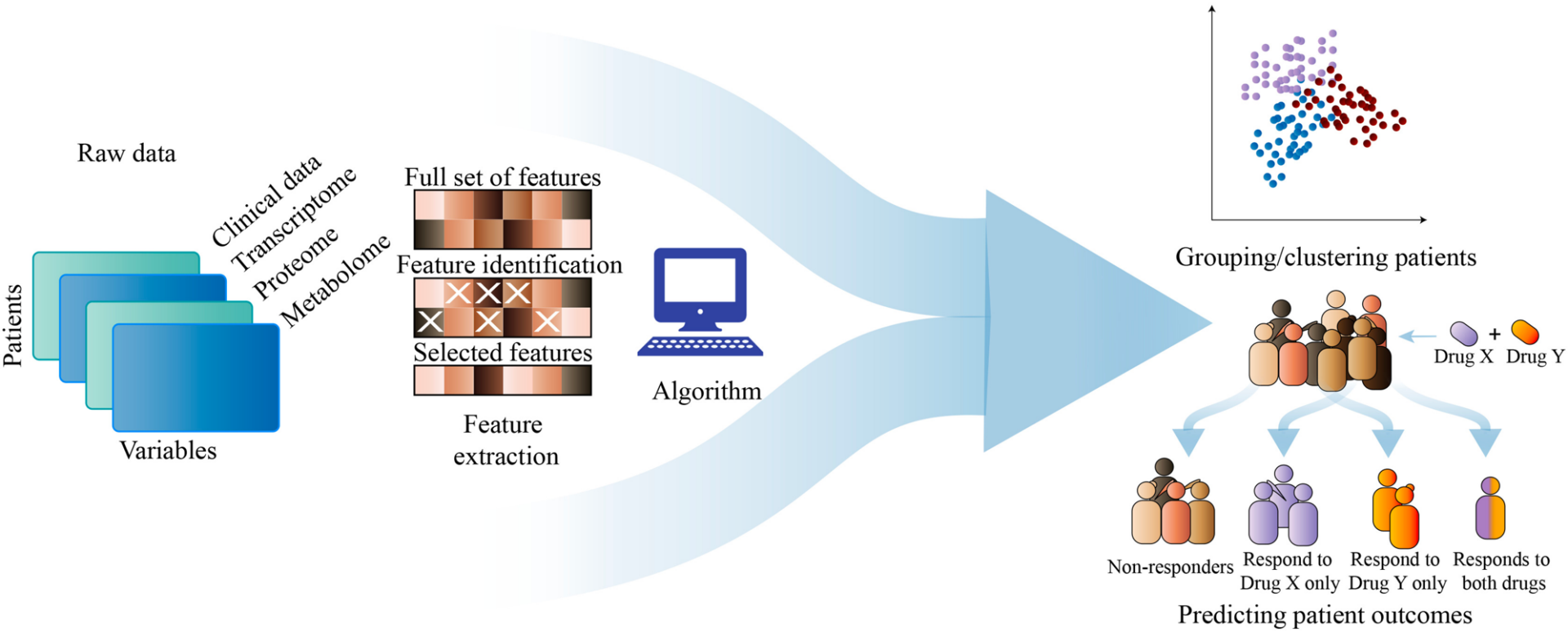
AI and ML

- Artificial Intelligence (AI) can be broadly defined as the science and engineering of making intelligent machines, especially intelligent computer programs
- Machine Learning (ML) is an AI technique that can be used to design and train software algorithms to learn from and act on data

https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

ML – Major Approaches

- Supervised learning
  - Algorithms are trained on labeled data, i.e. the desired output is known
- Unsupervised learning
  - Algorithms are trained on unlabeled data, i.e. the desired output is unknown
- Semisupervised learning, reinforcement learning, etc.



Toh TS, Dondelinger F, Wang D. Looking beyond the hype: Applied AI and machine learning in translational medicine. EBioMedicine. 2019;47:607-615.

Applications

- Drug discovery
  - Designing chemical compounds
  - Drug screening
- Imaging
  - Cell microscopy and histopathology
  - Radiology
- Genomic medicine
  - Biomarker discovery
  - Integrating different modalities of data

Toh TS, Dondelinger F, Wang D. Looking beyond the hype: Applied AI and machine learning in translational medicine. EBioMedicine. 2019;47:607-615.

Example Applications

Unsupervised hierarchical clustering (part of ACME analysis)

- Identified associations between *BRAF* mutant cell lines of the skin lineage being sensitive to the MEK inhibitör

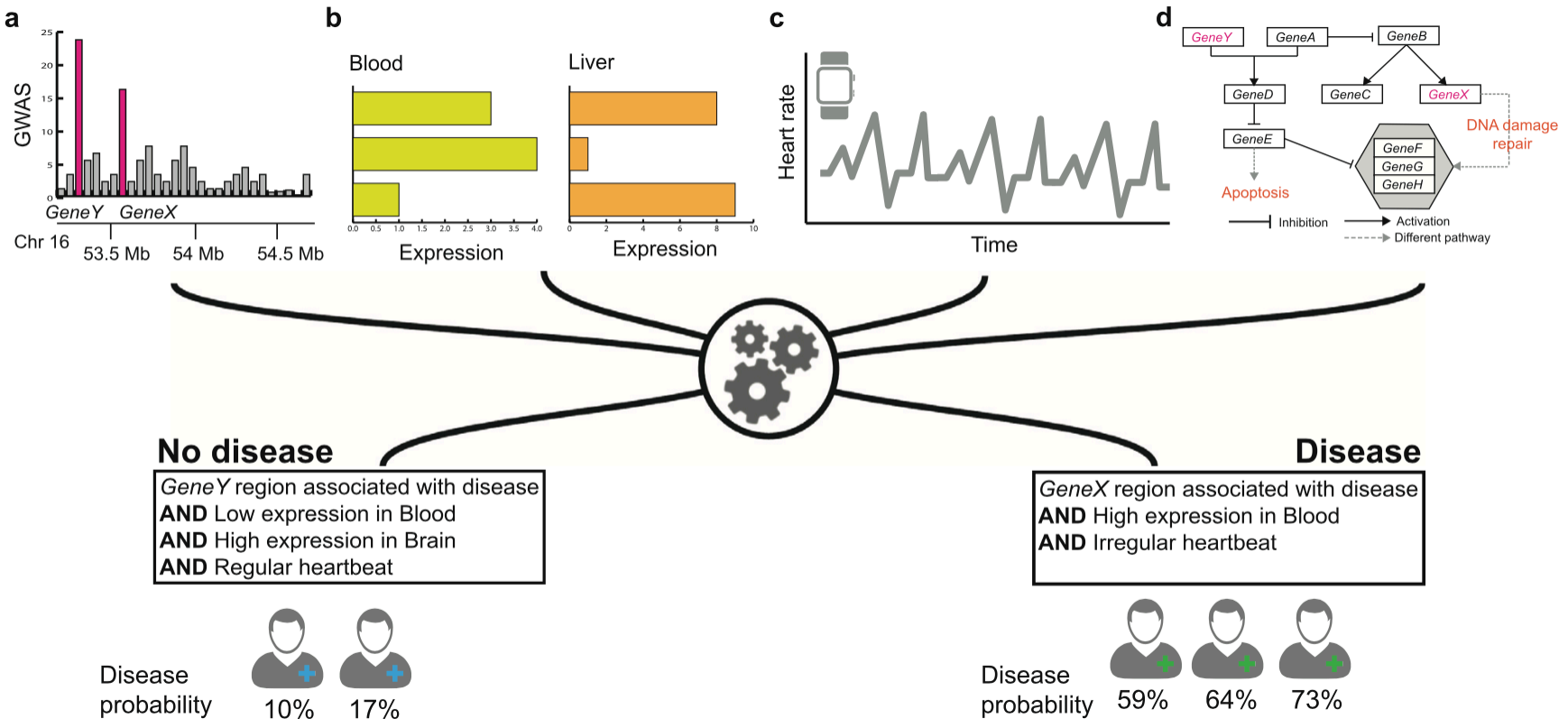
Seashore-ludlow B, Rees MG, Cheah JH, et al. Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. Cancer Discov. 2015;5(11):1210-23.

- Spectral clustering by SNF
  - Identification of new medulloblastoma subtypes

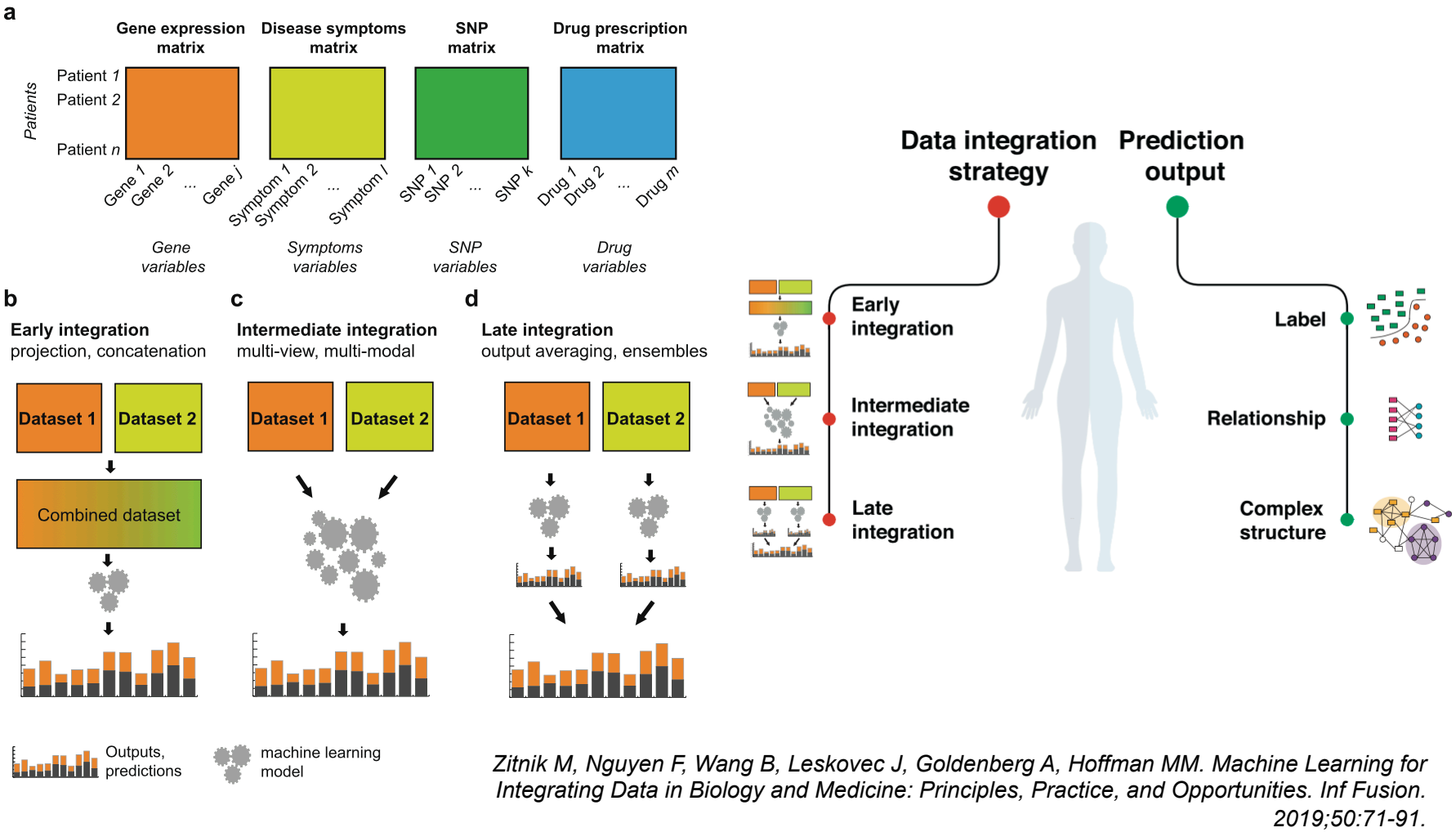
Cavalli FMG, Remke M, Rampasek L, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2017;31(6):737-754.e6.

- Elastic net regression
  - Identification of *BRAF* and *NRAS* mutations in cell lines, were among the top predictors of drug sensitivity for a MEK inhibitor

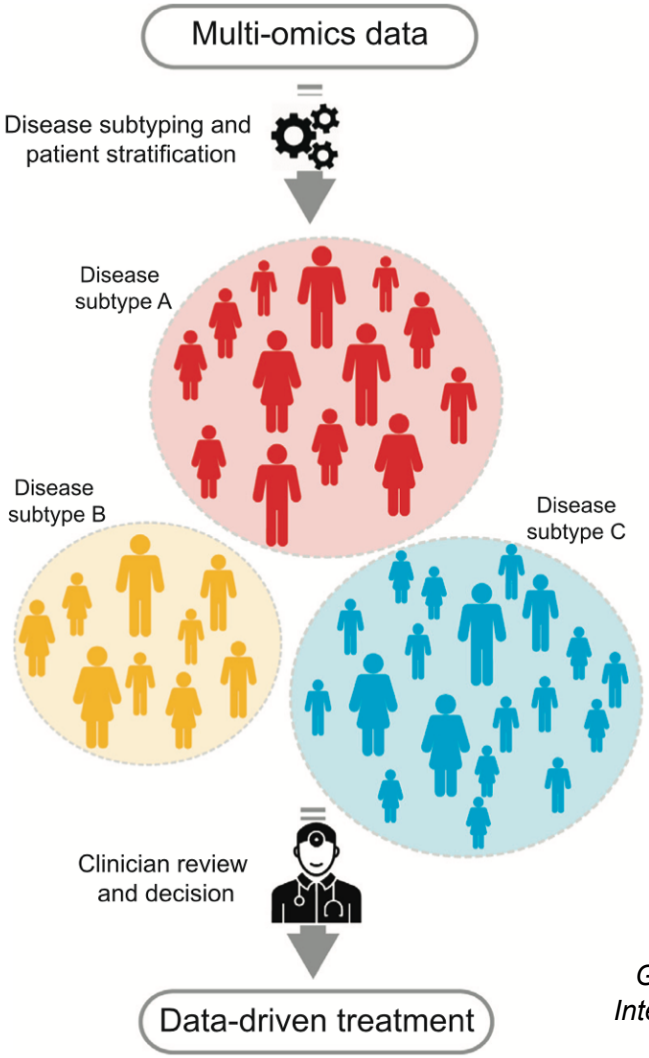
Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7.



Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine Learning for Integrating Data in Biology and Medicine: Principles, Practice, and Opportunities. Inf Fusion. 2019;50:71-91.

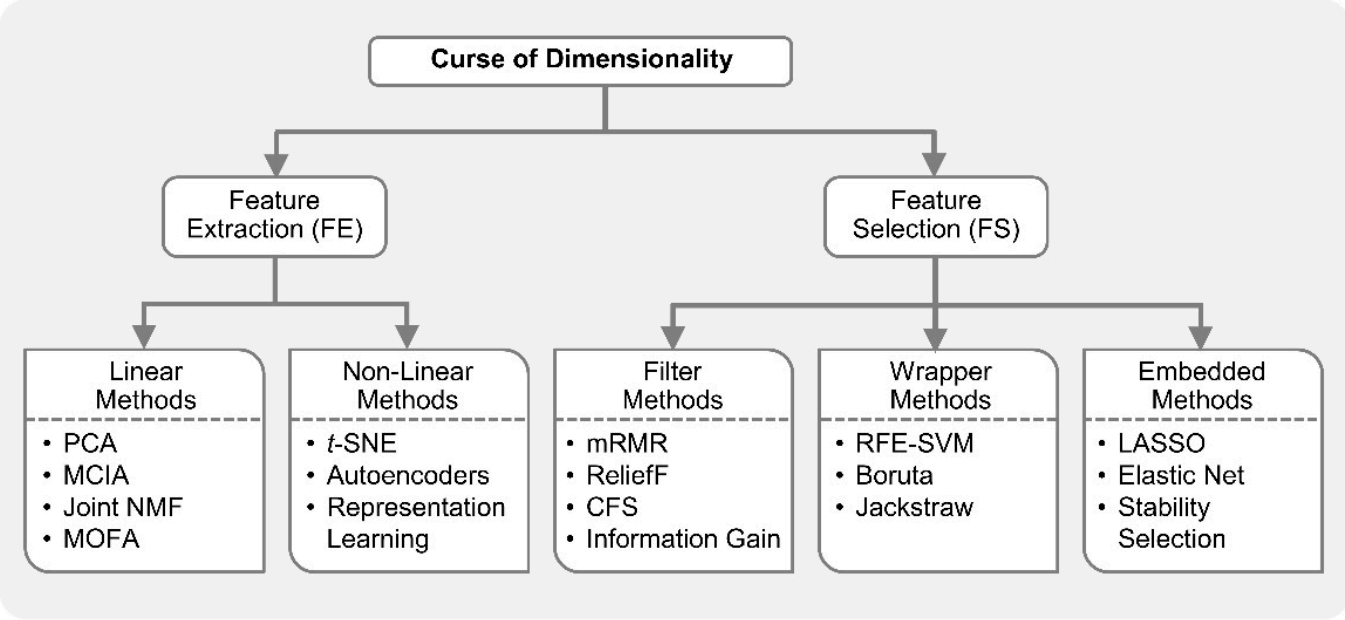


Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine Learning for Integrating Data in Biology and Medicine: Principles, Practice, and Opportunities. Inf Fusion. 2019;50:71-91.



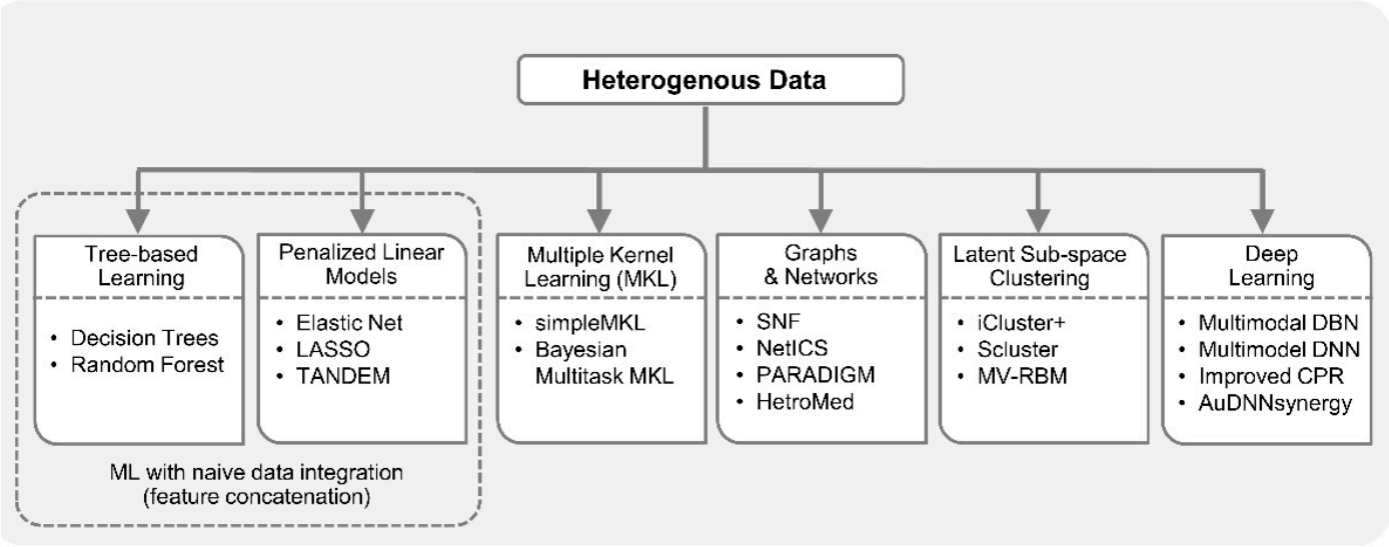
Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine Learning for Integrating Data in Biology and Medicine: Principles, Practice, and Opportunities. Inf Fusion. 2019;50:71-91.

Challenges



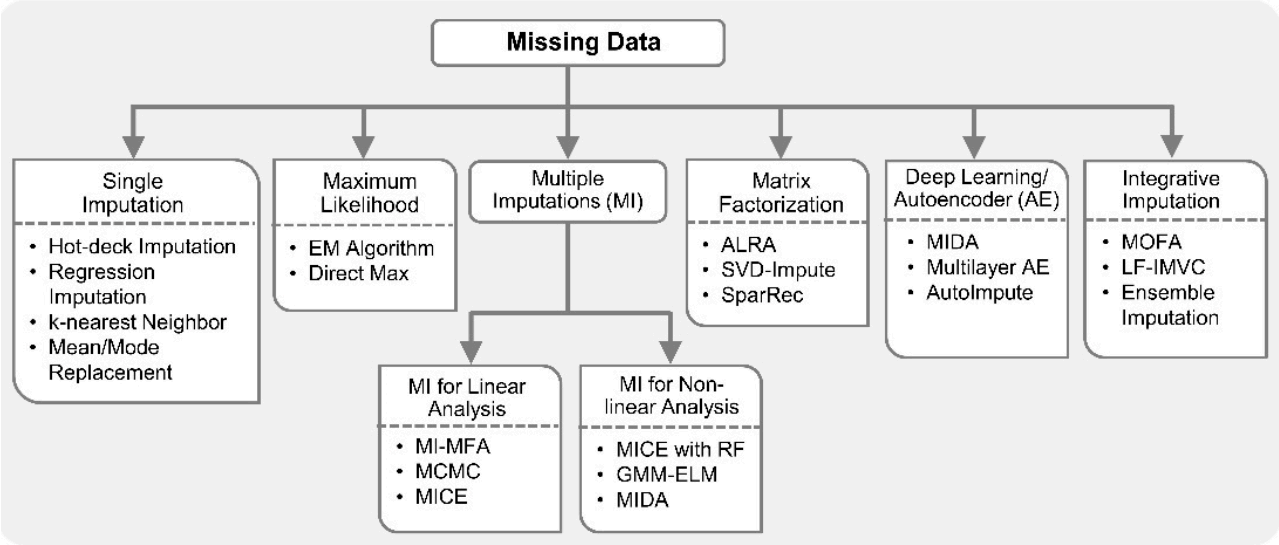
Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes (Basel). 2019;10(2)

Challenges



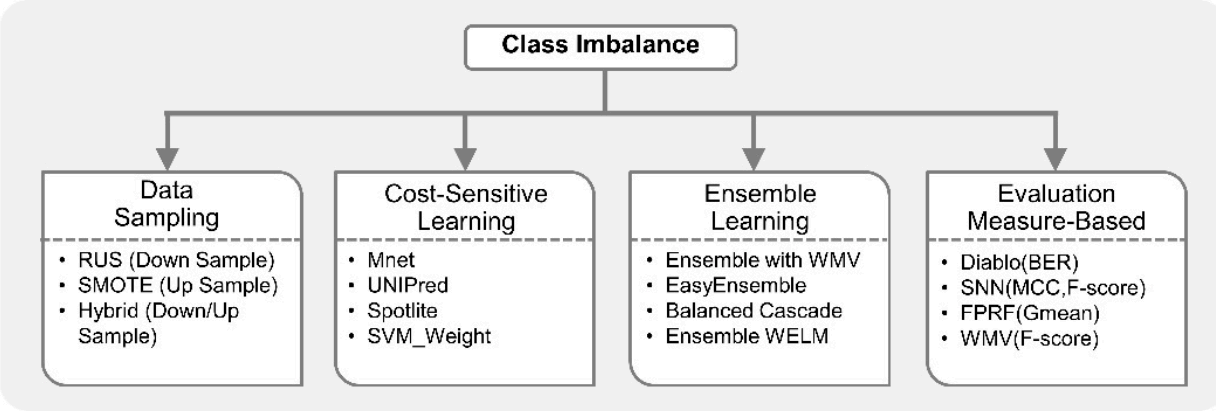
Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes (Basel). 2019;10(2)

Challenges



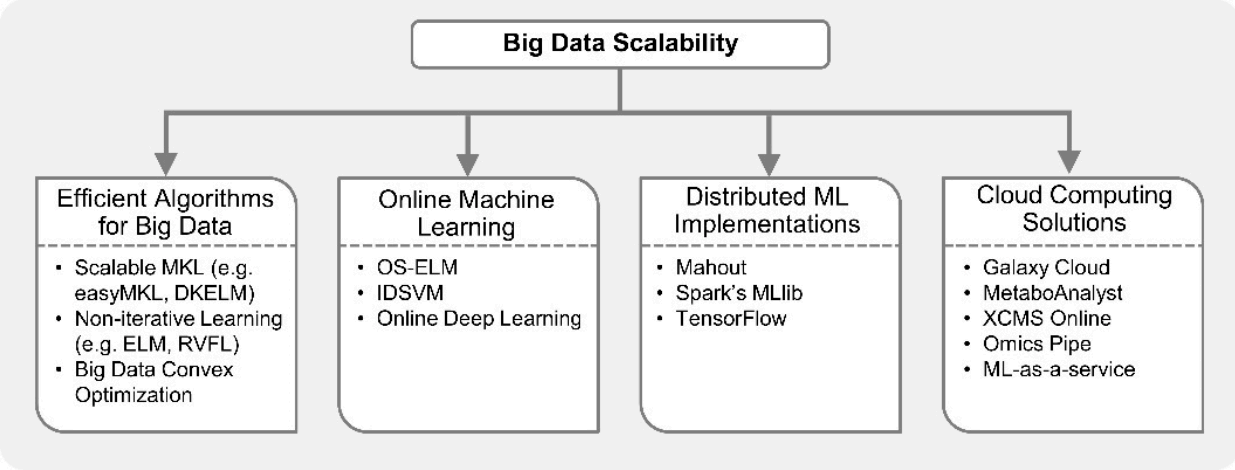
Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes (Basel). 2019;10(2)

Challenges



Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes (Basel). 2019;10(2)

Challenges

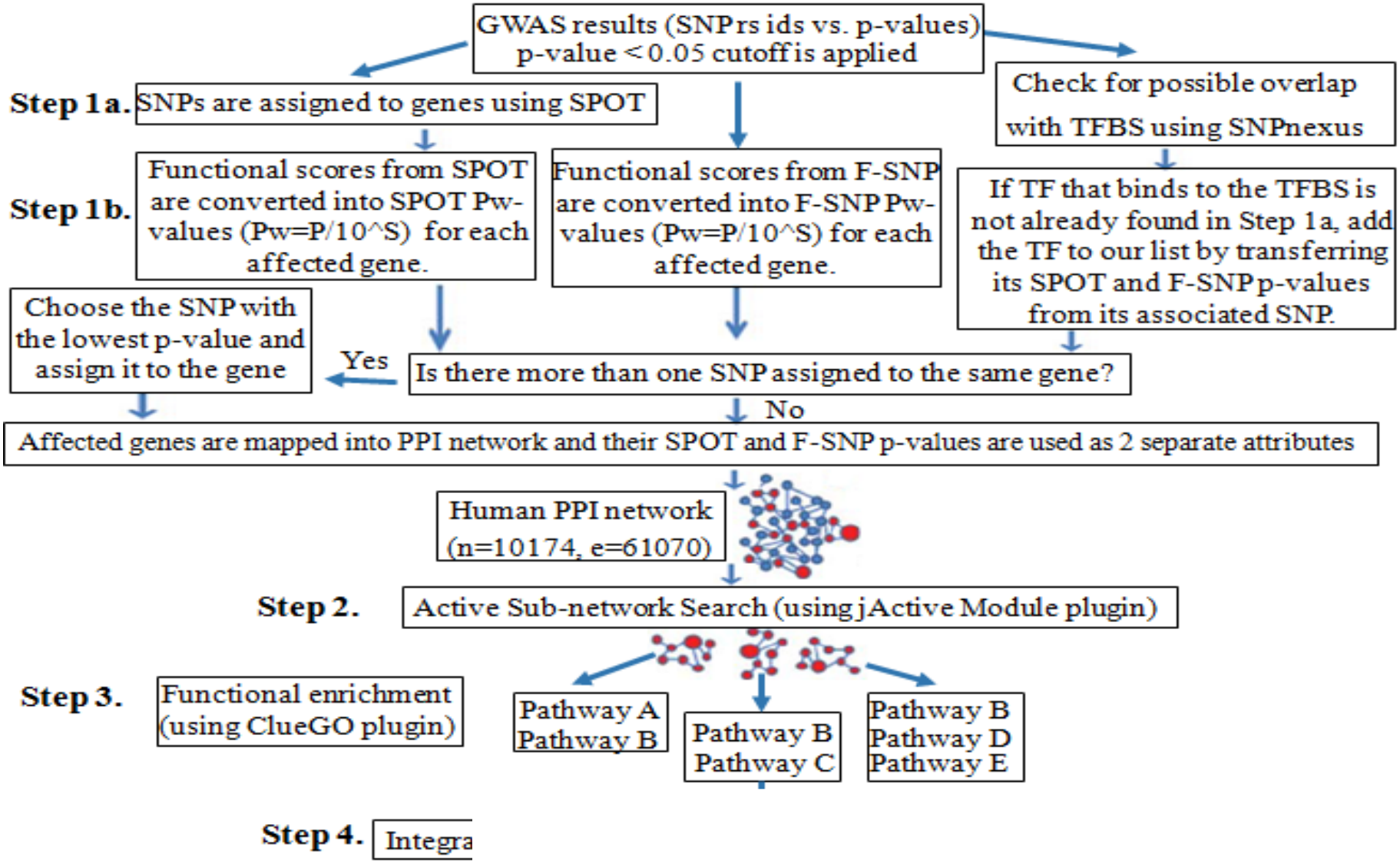


Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes (Basel). 2019;10(2)

Our Methodology

- NETWORK Based Integration of Omics Data

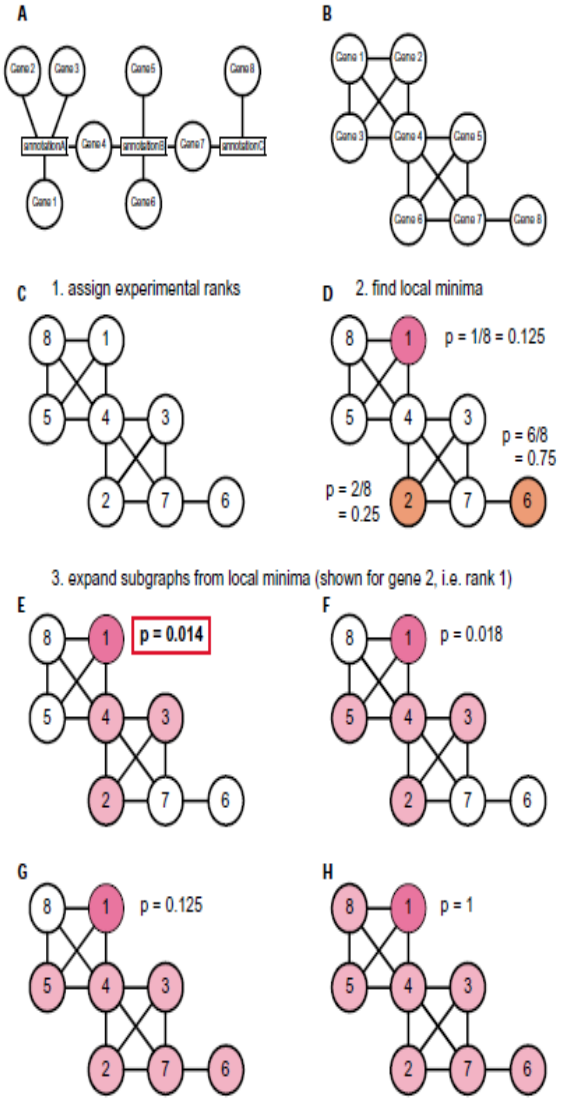
Our Methodology (PANOGA)



Active Subnetwork Search

- Breitling et al., 2004
  - mRNA expression data is used.
  - Significance ranks assigned to nodes.
  - Greedy search

$$p = \prod_{i=0}^{n-1} \binom{m-i}{N-i}$$



Partial Epilepsy Dataset

# of Cases	# of Controls	# of genotyped SNPs	Platform
3,445	6,935	528,745 SNPs	Illumina, Human610-Quadv1 genotyping chips

Table 5. Summary of Partial Epilepsy (PE)dataset (Kasperaviciute, et al., 2010).

- 1429 patients with epilepsies of unknown cause (classified as "cryptogenic"), 919 cases with mesial temporal lobe epilepsy with hippocampal sclerosis, 241 with cortical malformations and 222 patients with various tumors, other smaller subgroups such as trauma, stroke, perinatal insults, infections, etc.
- Cochran–Mantel–Haenszel test results were used as the genotypic p-values of the identified SNPs.
- Using $P<0.05$ cutoff:
  - 28,450 SNPs were included.

KEGG Term	p values	SNPs in GWAS	SNP Targeted Genes	Previous Studies Showing Support	Wang et al. Study	OMIM	GWAS on PE	CNV Study on Epilepsy	Epi GAD	Rogic et al. Study
Complement and coagulation cascades	2,16E-25	34	12	(Aronica, et al., 2008; Okamoto, et al., 2010)	-	Y	-	-	-	Y
Cell cycle	1,03E-24	24	14	(Aronica, et al., 2008; Jimenez-Mateos, et al., 2008; Limviphuvadh, et al., 2010)	-	Y	-	-	-	Y
Focal adhesion	7,10E-23	97	20	(Brockschmidt, et al., 2012)	Y	Y	Y	-	-	Y
ECM-receptor interaction	1,62E-22	62	14	(Aronica, et al., 2008)	Y	Y	-	-	-	Y
Jak-STAT signaling pathway	1,16E-21	24	16	(Jimenez-Mateos, et al., 2008; Okamoto, et al., 2010)	Y	Y	-	-	-	Y
MAPK signaling pathway	2,32E-19	73	23	(Jimenez-Mateos, et al., 2008; Okamoto, et al., 2010; Zhou, et al., 2011)	Y	Y	Y	-	Y	Y
Proteasome	1,15E-18	11	4	(Lauren, et al., 2010)	-	-	-	-	-	-
Ribosome	1,57E-18	2	2	(Lauren, et al., 2010)	-	-	-	-	-	Y
Calcium signaling pathway	5,73E-18	154	22	(Jimenez-Mateos, et al., 2008; Limviphuvadh, et al., 2010; Okamoto, et al., 2010; Zhou, et al., 2011)	Y	Y	Y	Y	Y	Y
Regulation of actin cytoskeleton	9,23E-18	88	19		Y	Y	-	Y	-	Y
Adherens junction	1,01E-17	79	13		-	-	Y	-	-	Y
Pathways in cancer	3,94E-17	112	22		Y	Y	Y	-	-	Y
Gap junction	6,32E-17	147	18	(Lauren, et al., 2010)	Y	Y	Y	-	-	Y
Apoptosis	3,72E-16	37	13	(Jimenez-Mateos, et al., 2008)	Y	Y	-	-	-	Y
Long-term depression	2,90E-15	151	15	(Lauren, et al., 2010)	Y	Y	Y	Y	Y	Y
Axon guidance	4,01E-15	59	12	(Jimenez-Mateos, et al., 2008; Limviphuvadh, et al., 2010)	-	-	-	-	-	Y
Fc gamma R-mediated phagocytosis	2,22E-14	66	12		Y	Y	Y	Y	-	Y
Tight junction	2,82E-14	82	13		Y	Y	Y	-	-	Y
ErbB signaling pathway	4,04E-14	86	12		Y	Y	Y	-	-	Y
Wnt signaling pathway	6,28E-14	44	13	(Aronica, et al., 2008; Okamoto, et al., 2010)	Y	Y	Y	-	-	Y

Table 6. Comparison of the top 20 SNP-targeted pathways with the pathways of the known genes, as associated to partial epilepsy.

Intracranial Aneurysm Dataset

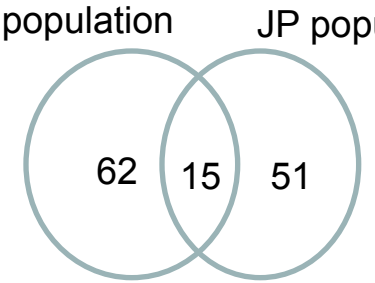
Population	# of Cases	# of Controls	# of genotyped SNPs	Platform
European	2,780	12,515	832,000	Illumina
Japanese	1,069	904	312,712	Illumina,

Table 7. Summary of Intracranial Aneurysm (IA)dataset.

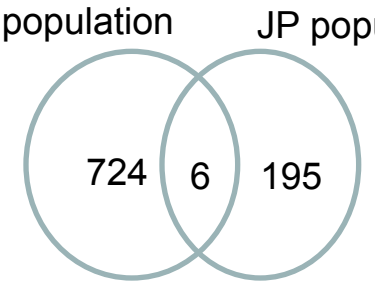
- In both datasets, each SNP's genotypic p-value of association is calculated via Cochran-Armitage trend test.
- Using P<0.05 cutoff:
  - 44,351 SNPs were included for EU population,
  - 14,034 SNPs were included for JP population.

KEGG Term	P-values EU	P-values JP	Rank EU	Rank JP	# of Associated SNPs in GWAS EU	# of Associated SNPs in GWAS JP	# of Common SNPs in GWAS	# of SNP Targeted Genes (STGs) EU	# of SNP Targeted Genes (STGs) JP	# of Com-mon STGs	% Common Genes in Both Populations EU	% Common Genes in Both Populations JP	Common SNPs in GWAS
MAPK signaling pathway *	3.53E-27	2.70E-18	1	8	133	43	1	14	18	2	14.29	11.11	rs791062
					5	18	1	11	10	2	18.18	20	rs744910
					6	20	3	15	9	5	33.33	55.56	rs2053423. rs1440375. rs744910
					0	15	0	6	4	0	0	0	
					7	45	1	21	14	5	23.81	35.71	rs4678167
					2	1	0	6	1	0	0	0	
					5	34	1	13	11	2	15.38	18.18	rs1561798
					5	13	0	8	4	1	12.5	25	
					2	36	1	18	14	1	5.556	7.143	rs4678167
					3	14	0	7	7	1	14.29	14.29	

# of SNP Targeted Genes in Top 10 Pathways



# of SNPs from GWAS in Top 10 Pathways



r both populations in IA. 7 out of the top 10 pathways are

related diseases in KEGG Disease Pathways Database.

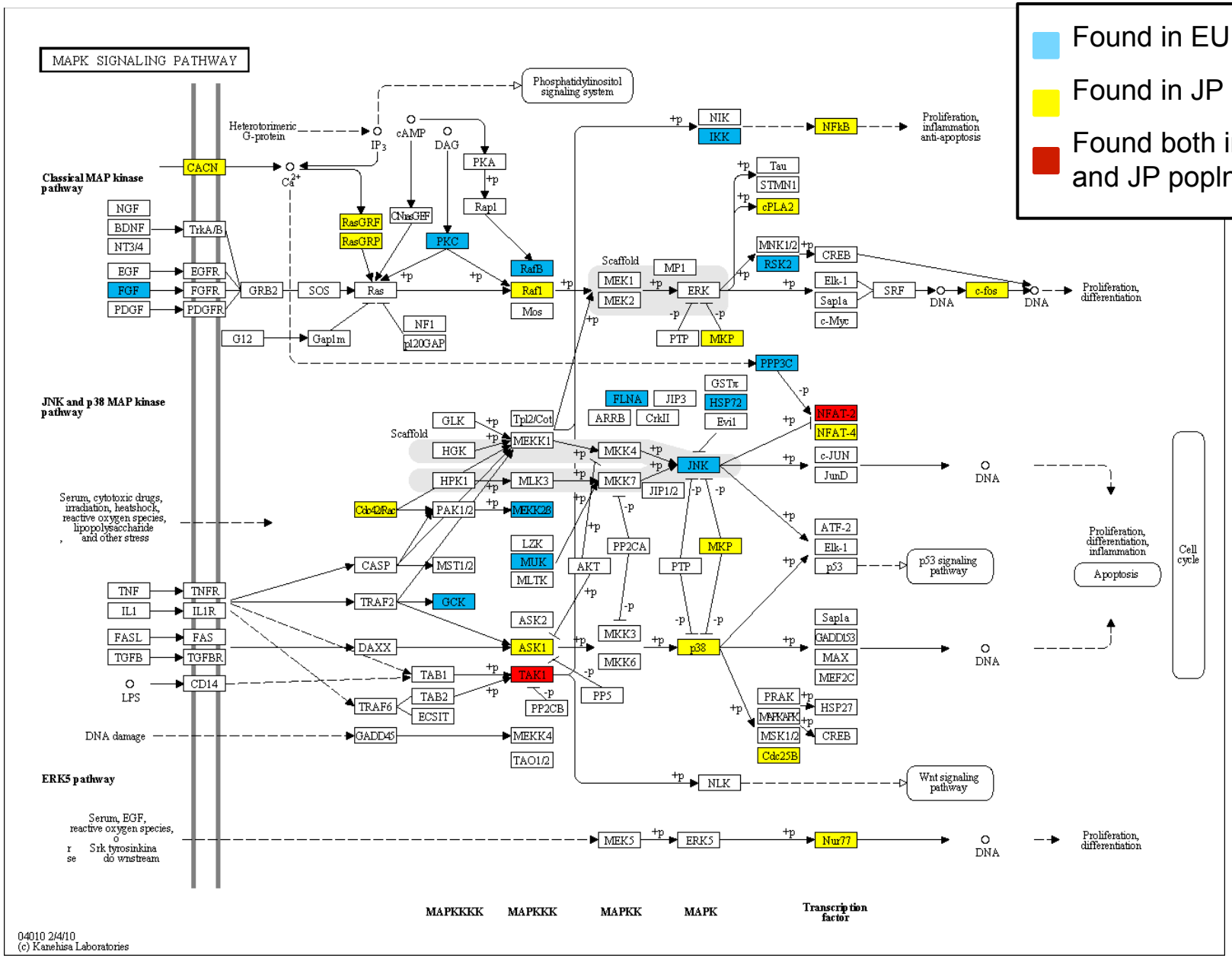


Figure 17. KEGG pathway map for MAPK signaling pathway. The set of genes shown in blue includes genes that are found for EU dataset; yellow includes genes that are found for JP dataset; red includes genes that are found both by EU and JP GWAS of IA.

Behcet's disease dataset

Population	# of Cases	# of Controls	# of genotyped SNPs	Platform
Turkish	1,215	1,278	311,459	Illumina, Infinium assay
Japanese	612	740	500,568	Affymetrix Gene Chip Human Mapping 500K

Table 10. Summary of Behcet's disease dataset.

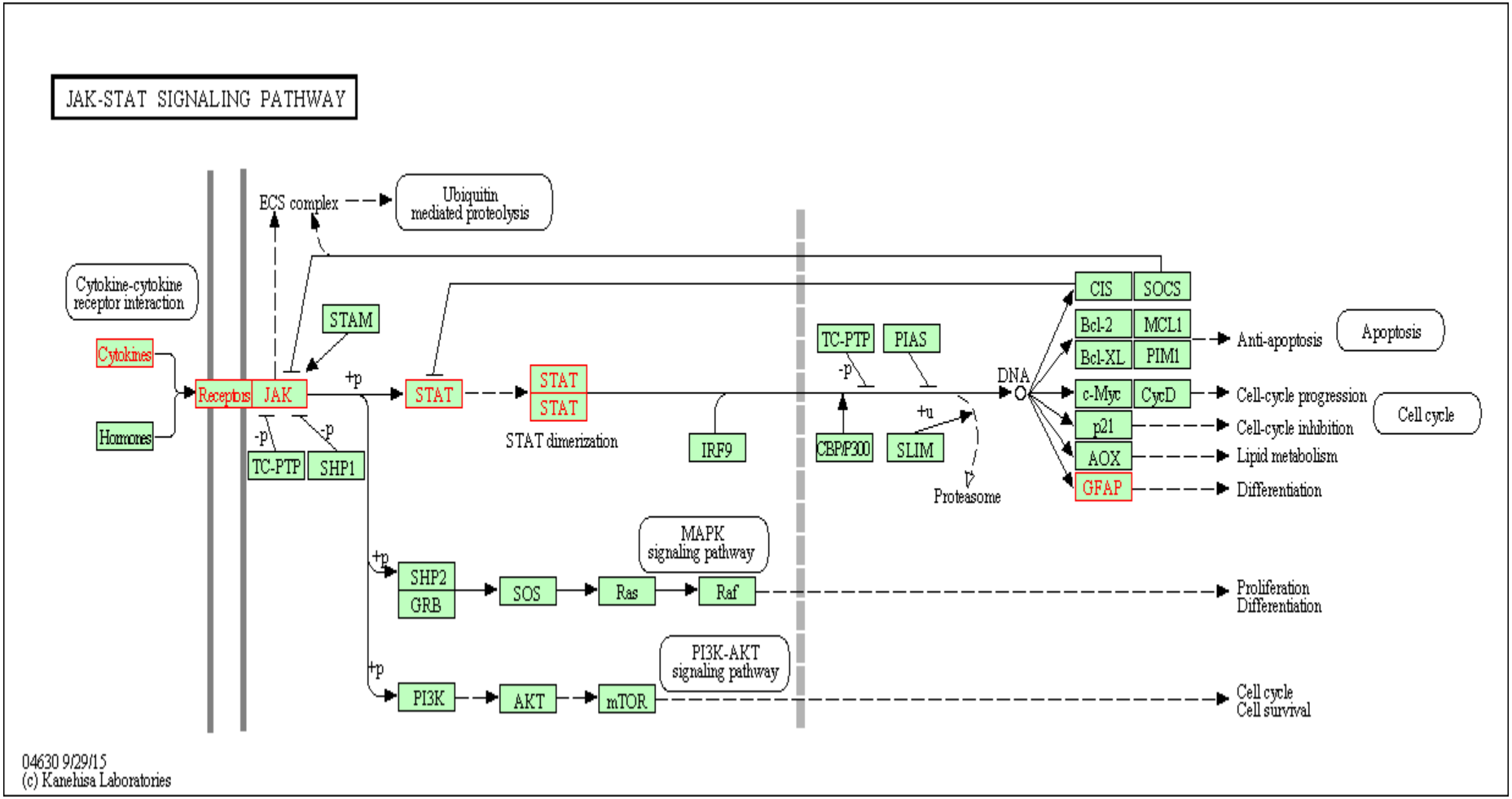
- In both datasets, each SNP's genotypic p-value of association is calculated via calculated via allelic chi-squared test.
- Using $P<0.05$ cutoff:
  - 18,479 SNPs were included for TR population,
  - 20,594 SNPs were included for JP population.

Common pathways in Turkish and Japanese Populations

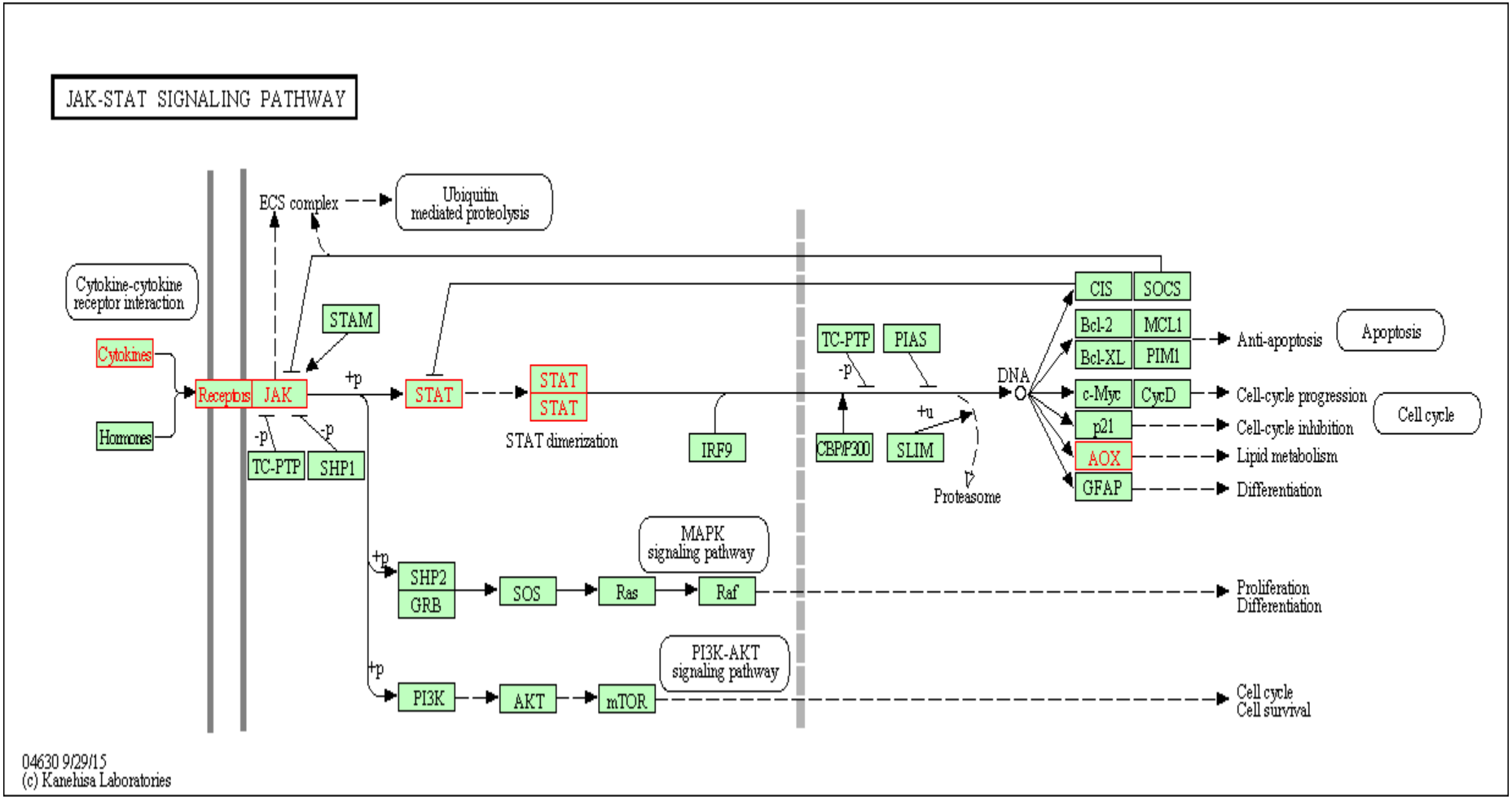
Antigen processing and presentation	Jak-STAT signaling pathway
Adipocytokine signaling pathway	MAPK signaling pathway
Aldosterone-regulated sodium reabsorption	Maturity onset diabetes of the young
Amoebiasis	NOD-like receptor signaling pathway
AMPK signaling pathway	Notch signaling pathway
Axon guidance	PPAR signaling pathway
cAMP signaling pathway	Prolactin signaling pathway
cGMP-PKG signaling pathway	Rap1 signaling pathway
Circadian rhythm	Ras signaling pathway
ErbB signaling pathway	Tight junction
Fc gamma R-mediated phagocytosis	Tuberculosis
Herpes simplex infection	Wnt signaling pathway
Inflammatory mediator regulation of TRP channels	

* Common pathways (25) in first 40 pathways of each population

Highest scoring Jak-STAT path in Turkish population

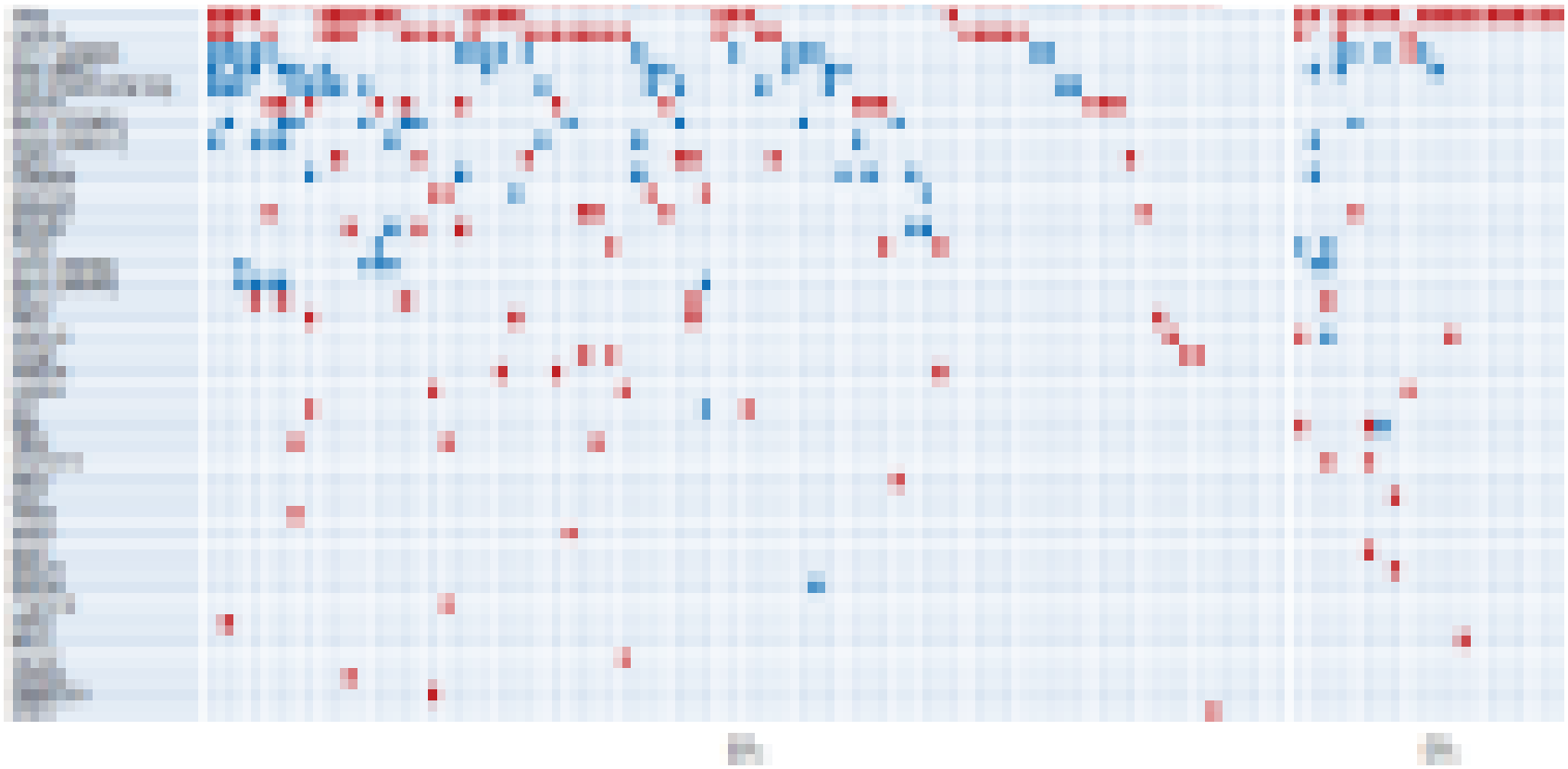


Highest scoring Jak-STAT path in Japanese population



Frequent cancers include high number of very rare genomic segments

- Somatic mutation
- Copy Number Variation



(whole genome sequencing breast cancers)

Stephens, Nature, 2012

Identification of Cancer drivers

- Identification of individualized driver mechanisms that lead to tumour specific cancer progression can improve patient's outcome

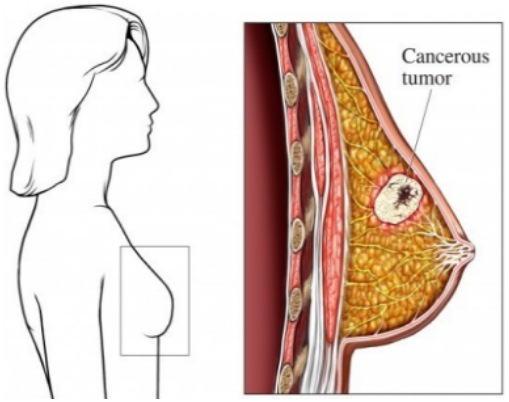
- Goal: Identification of targetable driver mechanism

Precision Medicine

Concept: Identify the targets to be treated in each patient

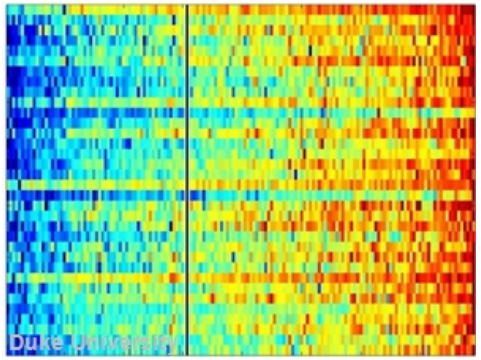
Clinical evidence

Therapy matched to genomic alteration



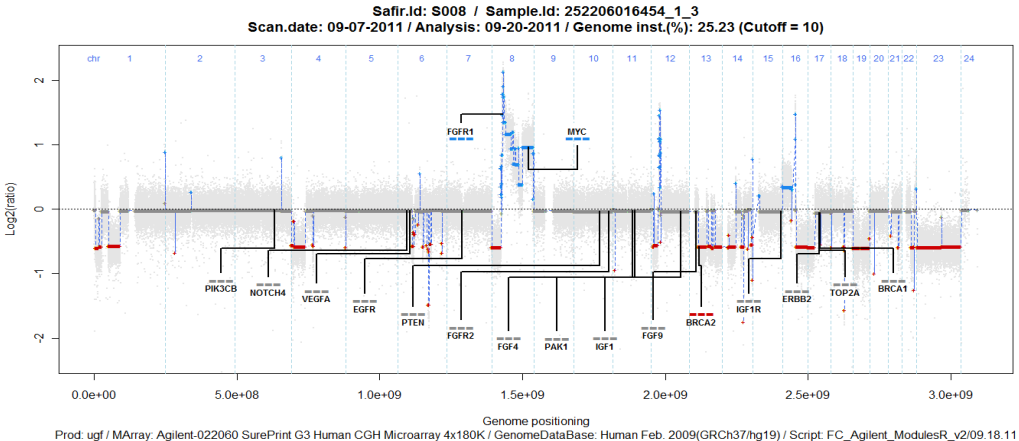
What is the optimal Biotechnology ?

Molecular analysis



What is the optimal Algorithm ?

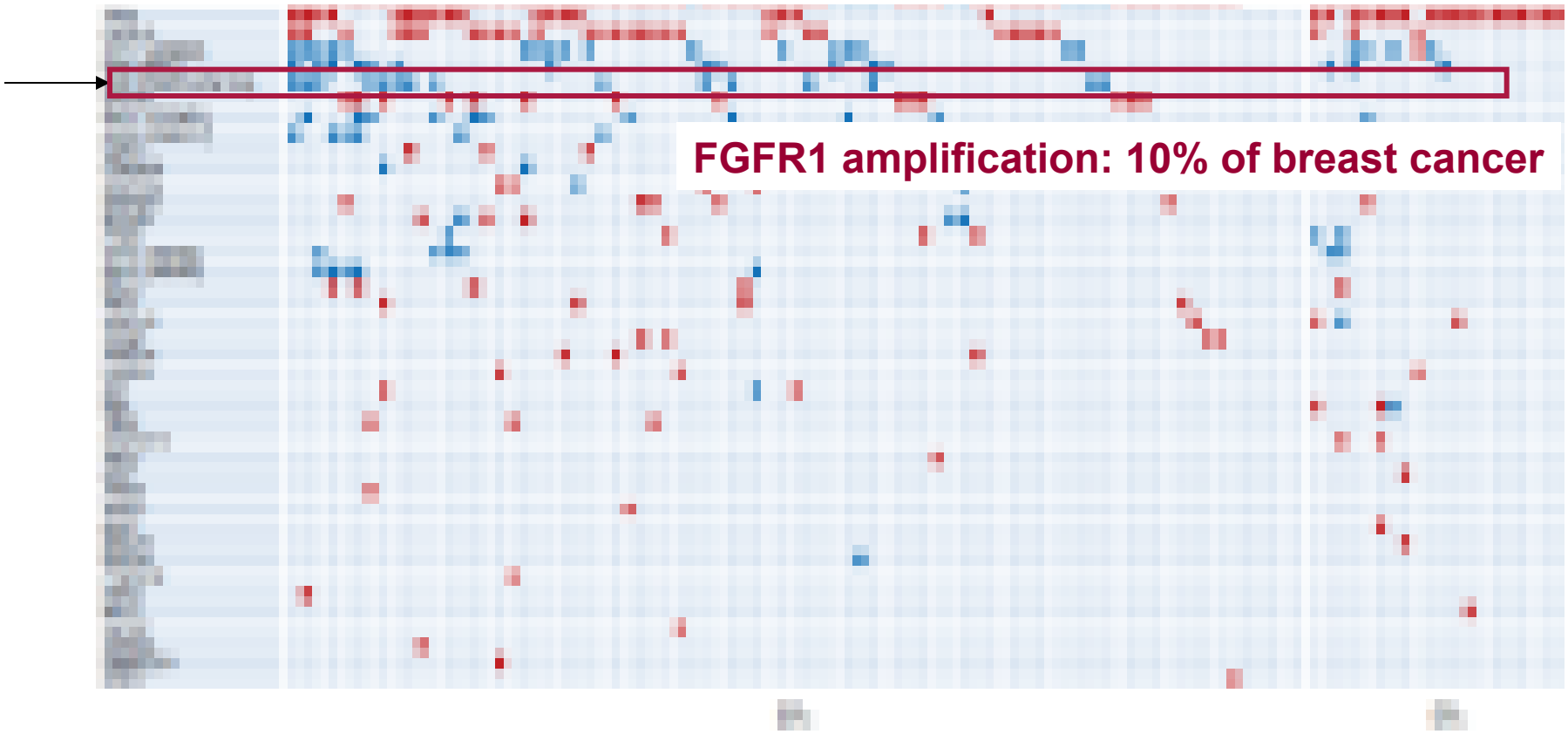
Target identification



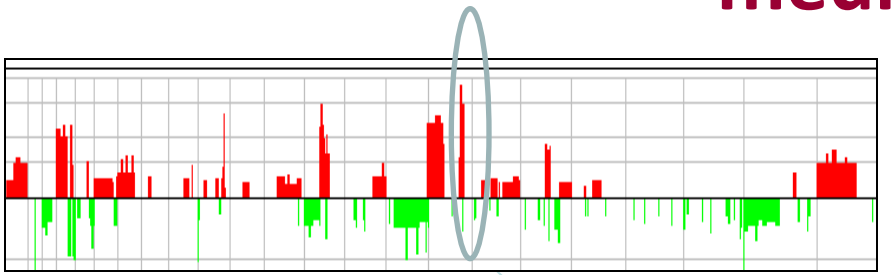
Andre, ESMO, 2012

Stratified medicine

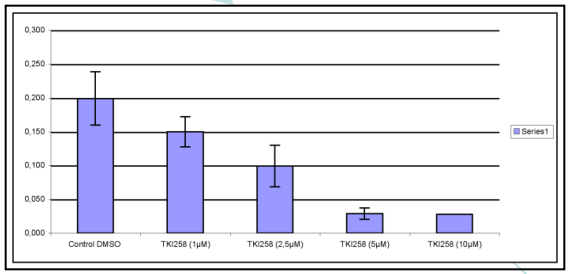
- Drug development or implementation in a strate defined by a molecular alteration



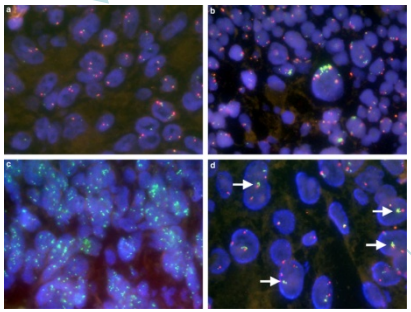
Translational research to feed stratified medicine



FGFR1: amplification in 10% BC

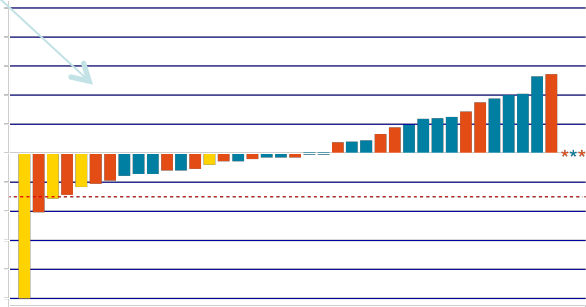


FGFR1 inhibitors present
higher sensitivity
on FGFR1-amplified CC

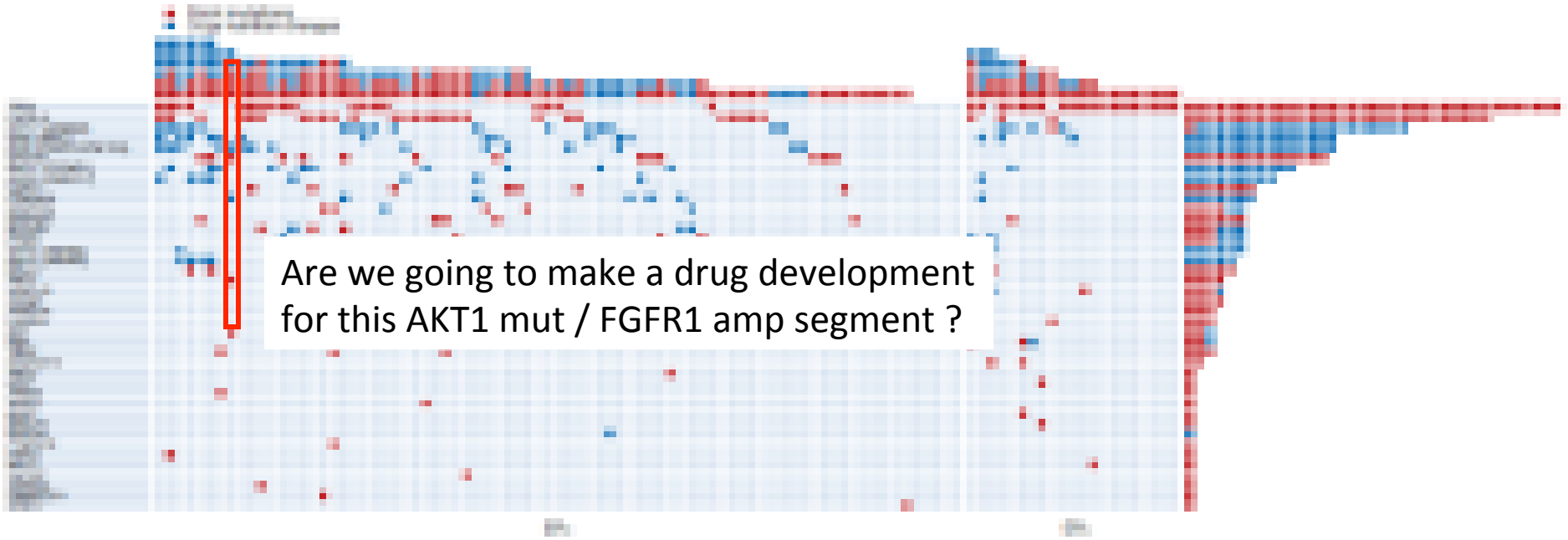


Set-up genomic test
(FISH)

Run phase II trial
Testing the FGFR1
Inh in patients with
FGFR1 amp BC



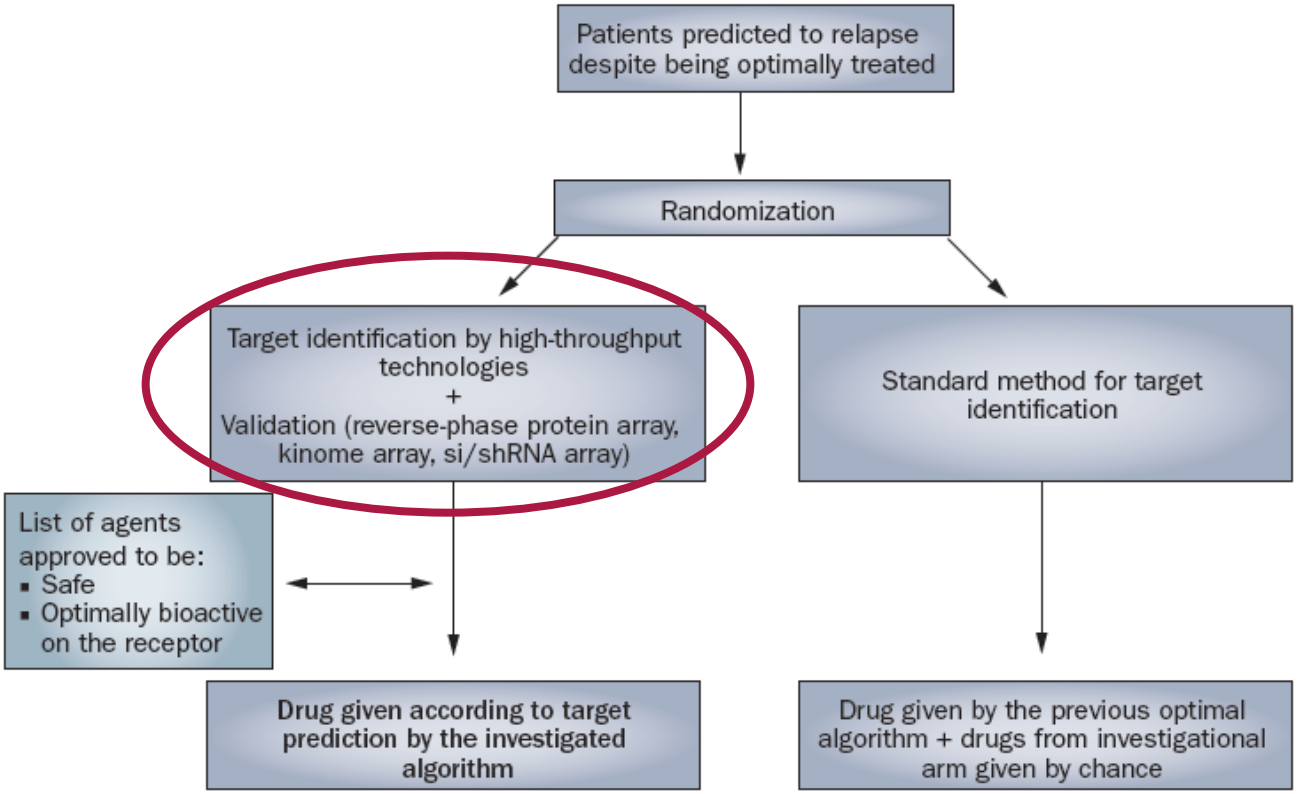
Evolution:
GENOMIC DISEASES ARE BECOMING TO RARE OR COMPLEX TO ALLOW DRUG DEVELOPMENT IN GENOMIC SEGMENTS



How to move forward ?

Stephens, Nature, 2012

Implications of Personalized Medicine



How to move there ???

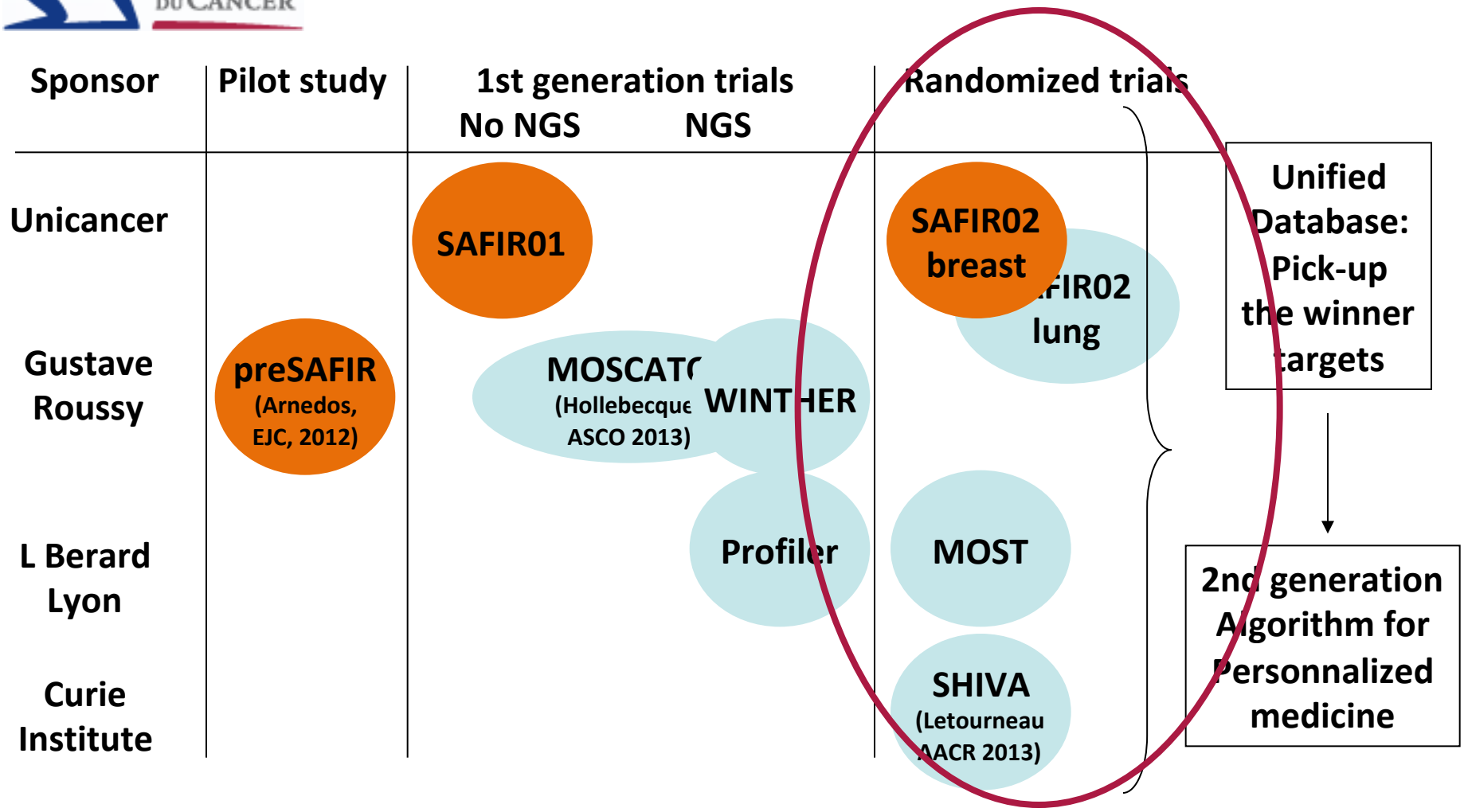
OPINION

Implications of personalized medicine —perspective from a cancer center

Thomas Tursz, Fabrice Andre, Vladimir Lazar, Ludovic Lacroix and Jean-Charles Soria

Tursz, T. *et al. Nat. Rev. Clin. Oncol.* **8**, 177–183 (2011)

Ongoing molecular screening or personalized medicine programs in France



Overall : >2 000 planned patients (all tumor types), >800 already included

Breast Cancer: > 1 000 planned, >70 already treated

Goal: To generate optimal algorithm for individualized therapy

SAFIR01

- 423 patients were included, and biopsy samples were obtained from 407 (metastatic breast cancer was not found in four). CGH array and Sanger sequencing were feasible in 283 (67%) and 297 (70%) patients, respectively.
- A targetable genomic alteration was identified in 195 (46%) patients, most frequently in PIK3CA (74 [25%] of 297 identified genomic alterations), CCND1 (53 [19%]), and FGFR1 (36 [13%]). 117 (39%) of 297 patients with rare genomic alterations (<5% of the general population), including AKT1 mutations, and EGFR, MDM2, FGFR2, AKT2, IGF1R, and MET high-level amplifications.
- Therapy could be personalised in 55 (13%) of 423 patients. Of the 43 patients who were assessable and received targeted therapy, four (9%) had an objective response, and nine others (21%) had stable disease for more than 16 weeks.
- Serious (grade 3 or higher) adverse events related to biopsy were reported in four (1%) of enrolled patients, including pneumothorax (grade 3, one patient), pain (grade 3, one patient), haematoma (grade 3, one patient), and haemorrhagic shock (grade 3, one patient).

A Protocol to Determine Somatic Modifications

- Exome Sequencing of tumour sample and control sample(Blood)
- Identification of somatic alterations in the tumour

Driver mutations

Copy Number Variations (CNV)

SNPs

Chr	Pos	Ref -> Alt Genome Protein Effect	Gene	dbSNP	CGC* Tumor Type	DrugBank
2	209113112	C -> T R -> H Missense	IDH1	rs121913500	Glioblastoma	-
17	7577545	T -> C M -> V Missense	TP53	rs483352695 rs397516437	Glioma	Acetylsalicylic acid

* Cancer Gene Census

SNPs

Chr	Pos	Ref -> Alt Genome Protein Effect	Gene	dbSNP	CGC* Tumor Type	DrugBank
2	209113112	C -> T R -> H Missense	IDH1	rs121913500	Glioblastoma	-
17	7577545	T -> C M -> V Missense	TP53	rs483352695 rs397516437	Glioma	Acetylsalicylic acid

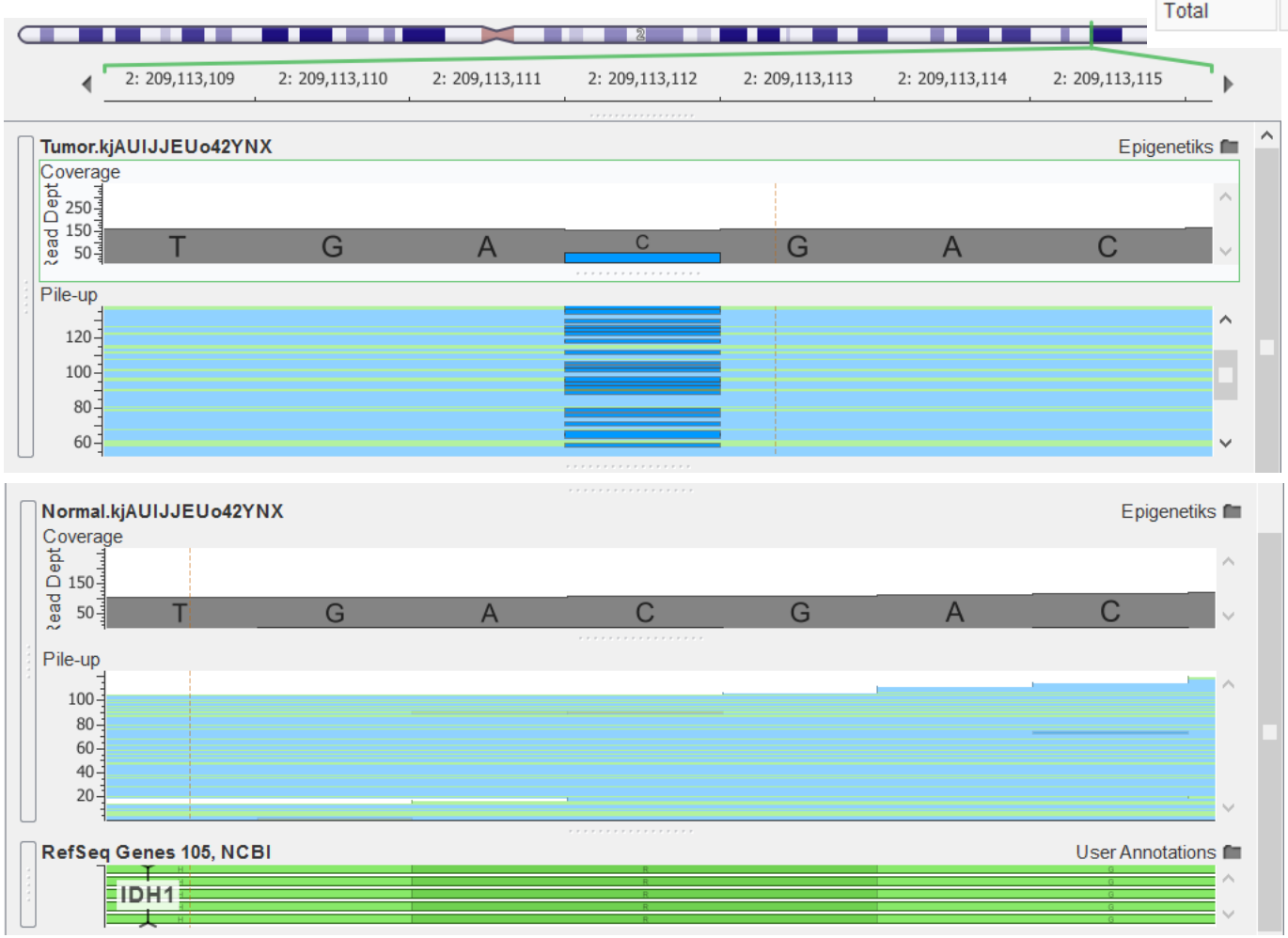
* Cancer Gene Census

SNPs

Chr2: 209,113,112

Matches / Mismatches / Deletions

Type	Base	Count	% of Total	Mean Quality
(match)	C	99	66.4	30.8
(mismatch)	T	50	33.6	32.3
Total		149	100	31.3



Copy Number Variation

Chr	Start	End	Normal Depth	Tumor Depth	Log Ratio
8	29523990	29524007	20.6	4.3	-2.348

CNV Annotation

CNV type	Disease	Platform	Pubmed
Deletion	Medulloblastoma	SNP arrays	21979893
Loss	Glioblastoma multiforme	CGH	19960244
Loss	Glioblastoma multiforme	conventional CGH	21080181
Loss	Glioblastoma multiforme	aCGH	21080181
Loss	Medulloblastoma	CGH	16968546

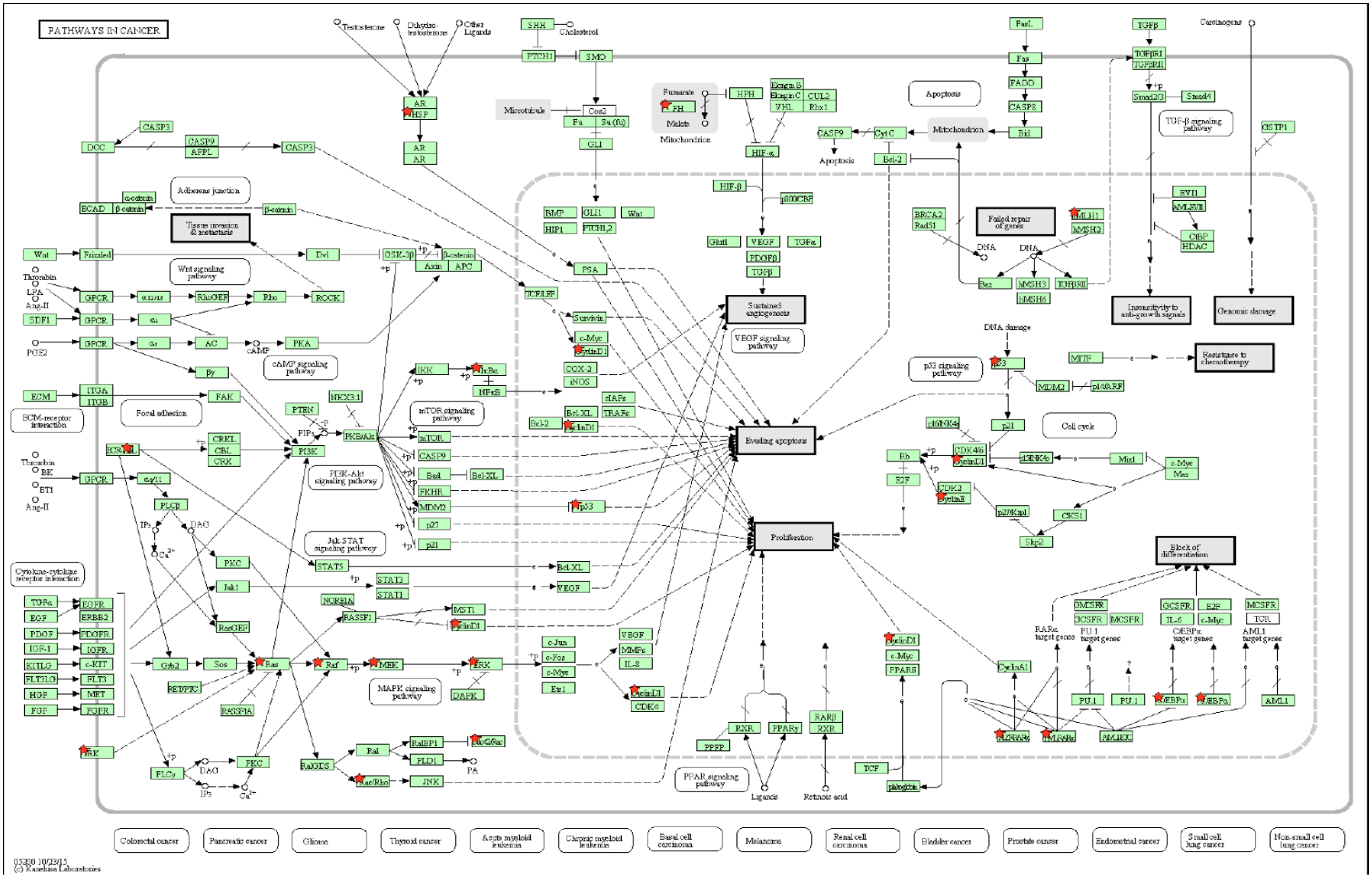
Scoring Algorithm

- Scoring system to identify major pathways leading to tumor progress
- Scoring System for targetable alterations in the tumor
- Scoring system for available drugs targeting most of the driver alterations

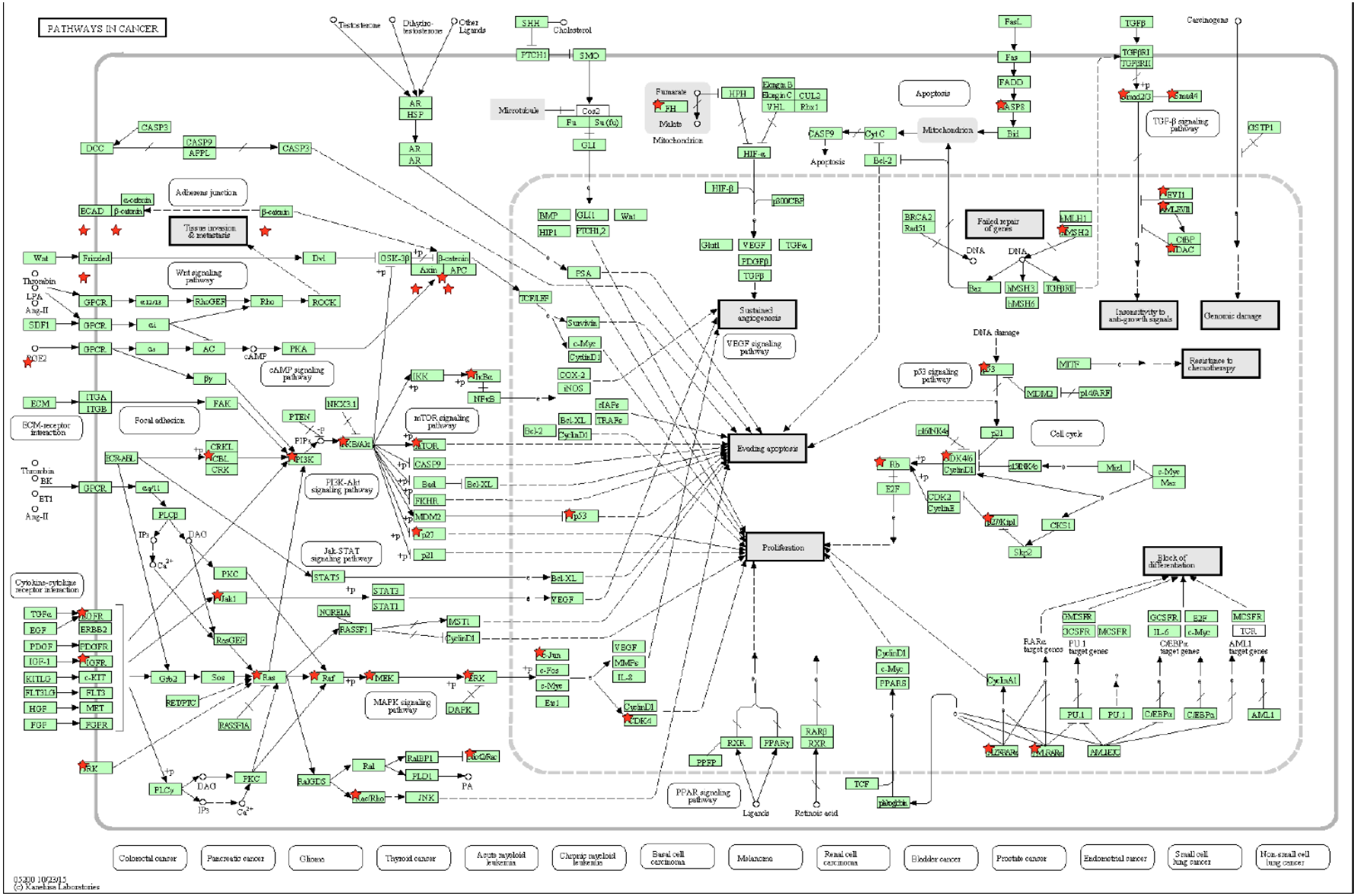
EXAMPLES of Exome Sequencing Data

- Patient 1 has CyclinD1 pathway over activated
- Patient 2 has Mtor pathway and CDK4 pathway activate
- Patient 3 has over amplification of Growth Factor receptors along with c-myc amplification
- Each has different driver mechanisms and requires different theraupeutical scheme

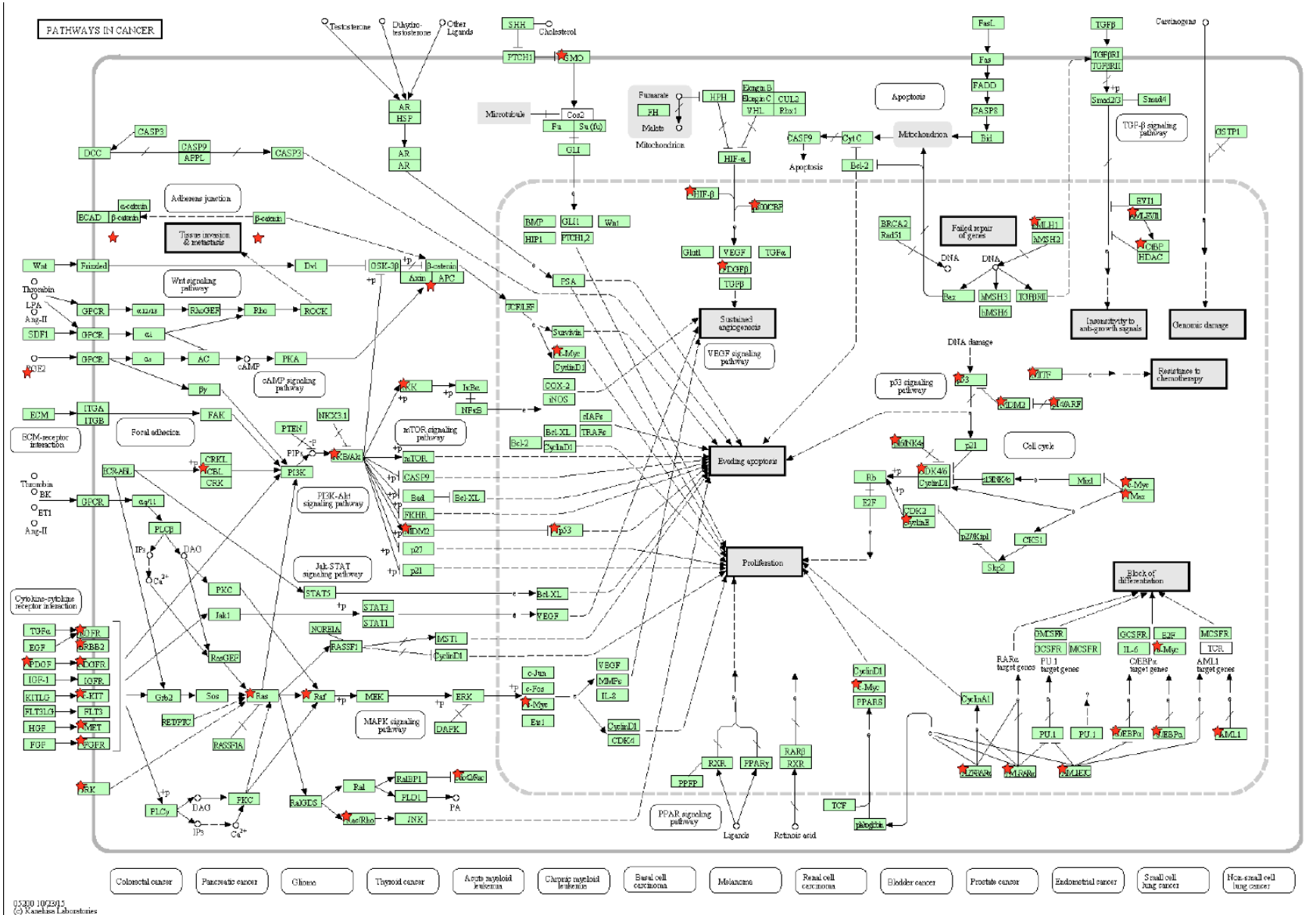
Patient 1


PATHWAYS IN CANCER
Colorectal cancer
Pancreatic cancer
Glioma
Thyroid cancer
Acute myeloid leukemia
Chronic myeloid leukemia
Basal cell carcinoma
Melanoma
Renal cell carcinoma
Bladder cancer
Prostate cancer
Endometrial cancer
Small cell lung cancer
Non-small cell lung cancer

Patient 2


PATHWAYS IN CANCER
Colorectal cancer
Pancreatic cancer
Glioma
Thyroid cancer
Acute myeloid leukemia
Chronic myeloid leukemia
Basal cell carcinoma
Melanoma
Renal cell carcinoma
Bladder cancer
Prostate cancer
Endometrial cancer
Small cell lung cancer
Non-small cell lung cancer
05200 10/23/15
(c) Kanehisa Laboratories

Patient 3


PATHWAYS IN CANCER
Proliferation
Evading apoptosis
Sustained angiogenesis
Invasion & metastasis
Insensitivity to anti-growth signals
Genomic damage
Resistance to chemotherapy
Cell cycle
Block of differentiation
Colorectal cancer
Pancreatic cancer
Glioma
Thyroid cancer
Acute myeloid leukemia
Chronic myeloid leukemia
Basal cell carcinoma
Melanoma
Renal cell carcinoma
Bladder cancer
Prostate cancer
Endometrial cancer
Small cell lung cancer
Non-small cell lung cancer

Genotyping for prevention Timoma (Sternum)Patient 4

- AMPD1 chr1 115236056_115236057

GA 192 snp rs17602729 Caa/Taa

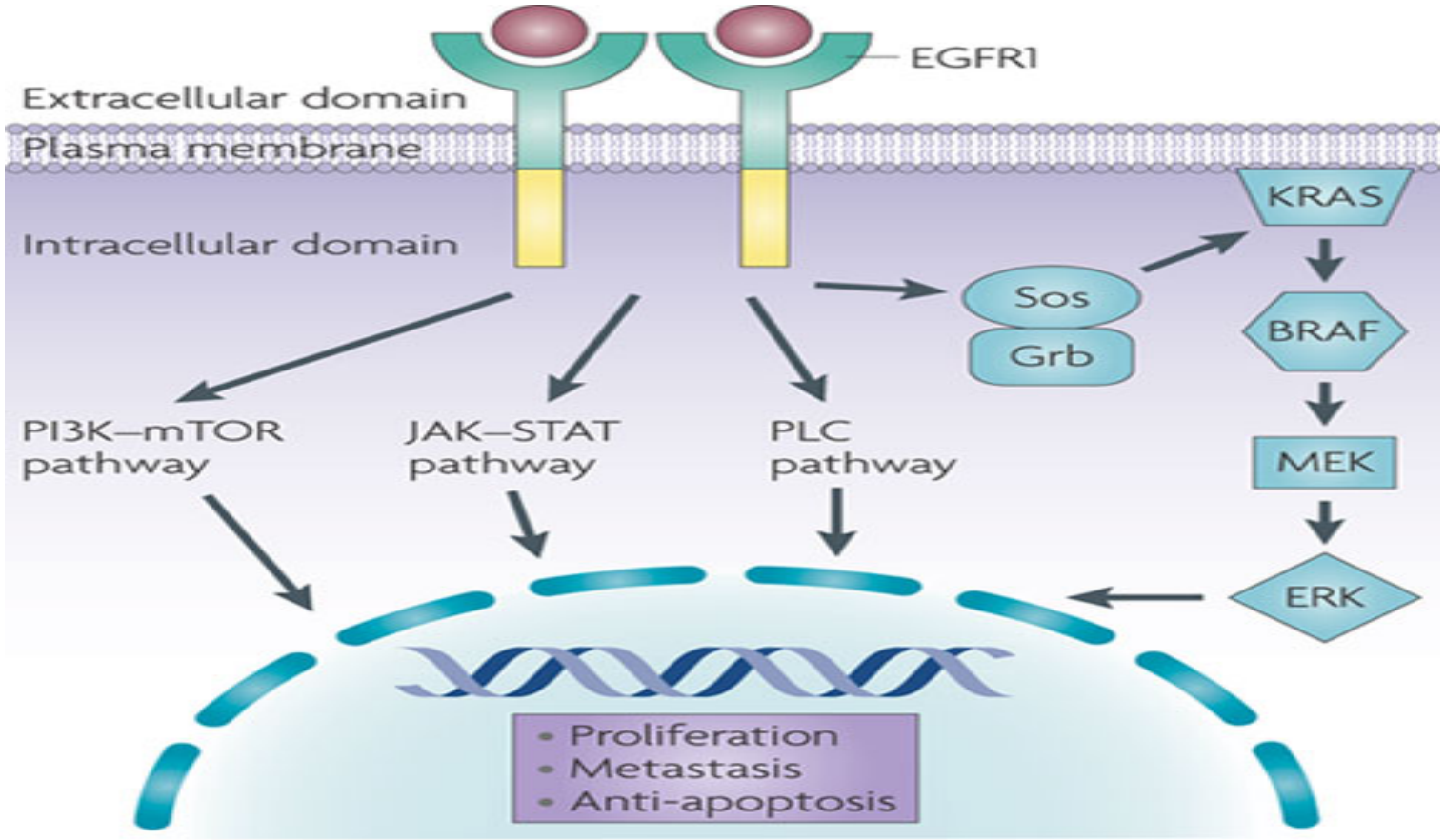
Q/* protein_coding stop_gain stop_gained HIGH pathogenic

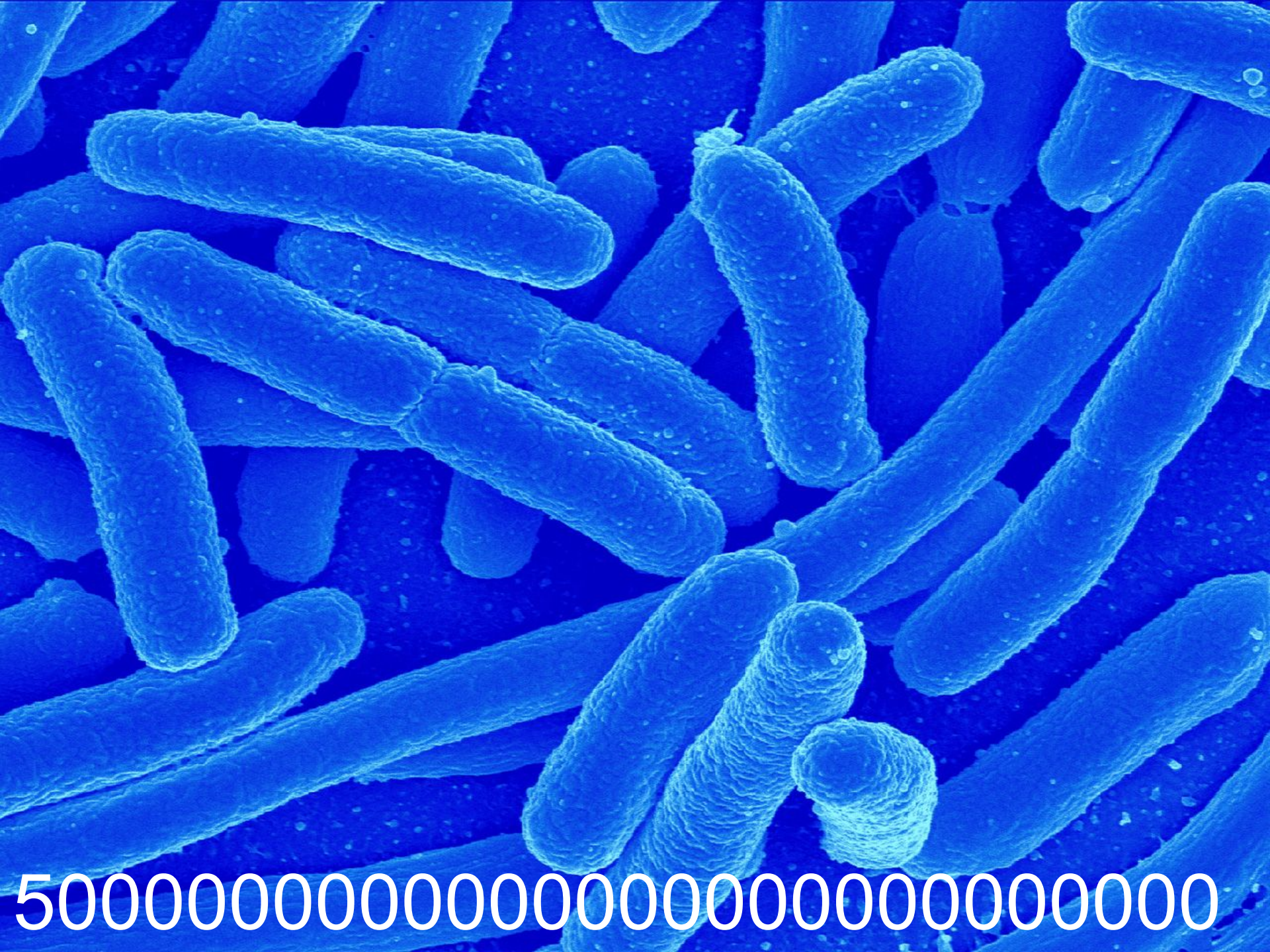
Muscle_AMP_deaminase_deficiency|

Myestenia Gravis

Target	Cancer	Variation type	Marker	Drug	Test
EGFR	Lung cancer	Mutation	Predict benefit to EGFR TKIs	Erlotinib	DNA
				Gefitinib	
ALK	Lung cancer	Rearrangement	Predict response to ALK inhibitors	Crizotinib	FISH
ROS	Lung cancer	Rearrangement	Predict response to TKIs	Crizotinib	FISH
RET	Lung cancer	Rearrangement	Predict response to TKIs	Vandetanib	FISH
BRAF	Melanoma	Mutation	Predict response to BRAF inhibitors	Vemurafenib	DNA
				Dabrafenib	
KRAS	Colorectal cancer	Mutation	Predict lack of response to anti-EGFR antibodies	Panitumumab	DNA
				Cetuximab	
HER2	Breast cancer	Amplification	Predict response to anti-HER2 antibodies	Trastuzumab	FISH, IHC
	Gastric cancer	Overexpression		Lapatinib	
				Pertuzumab	
KIT	GIST	Mutation	Predict response to c-Kit inhibitors	Imatinib	IHC
Estrogen receptor	Breast cancer	Overexpression	Predict response	Examestane	IHC
				Fulvestrant	
				Letrozole	
				Tamoxifen	
Progesterone receptor	Breast cancer	Overexpression	Predict response	Examestane	IHC
				Letrozole	

Personalized Treatment
Imatinib




5000000000000000000000000000

Our Microbiome Projects

- METASUB
- Breast Feeding vs. Formula Feeding (B. Infantes)
- Wellness Bioinformatics
- MS hastalarında fekal transplantasyon

İBS Fonksyonel bir hastalık mıdır? Türk Kohortlarında Prospektif, Kontrollü Mikrobiyota Çalışması

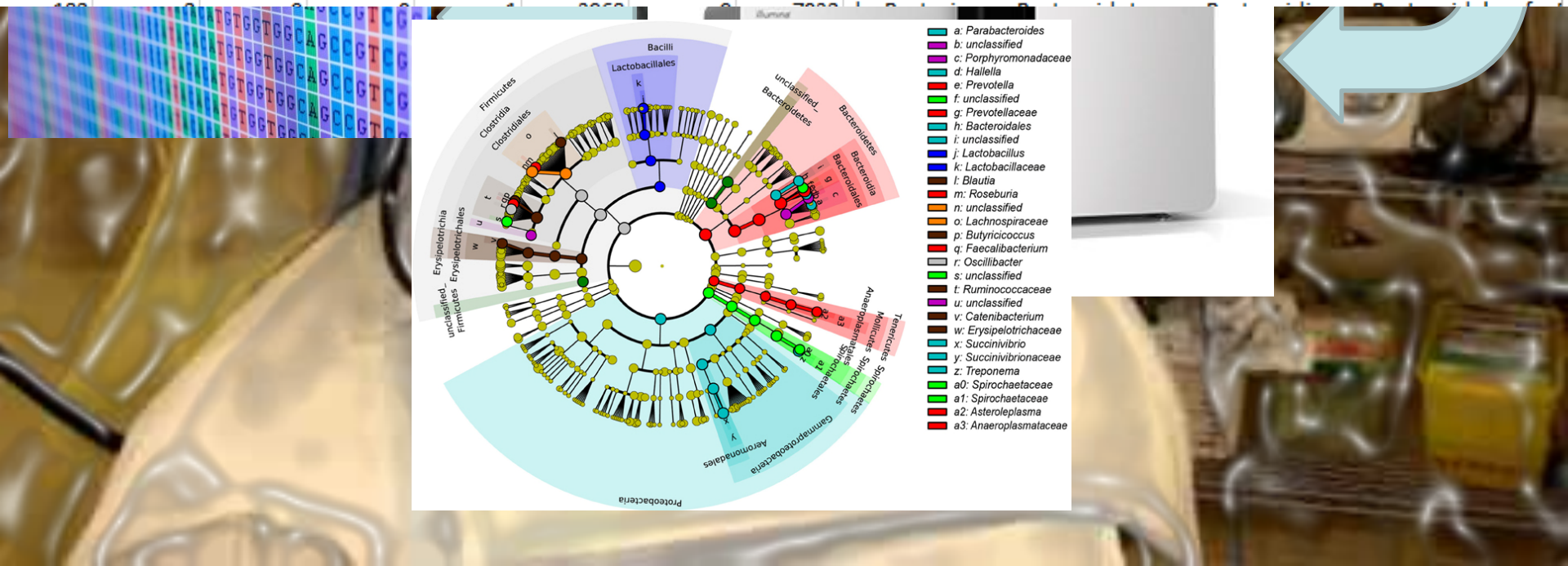
Munkhtsetseg Banzragch, Orhan Özcan, Osman Uğur Sezerman, Sinem Öktem, Özgür Kurt, **Nurdan Tözün**

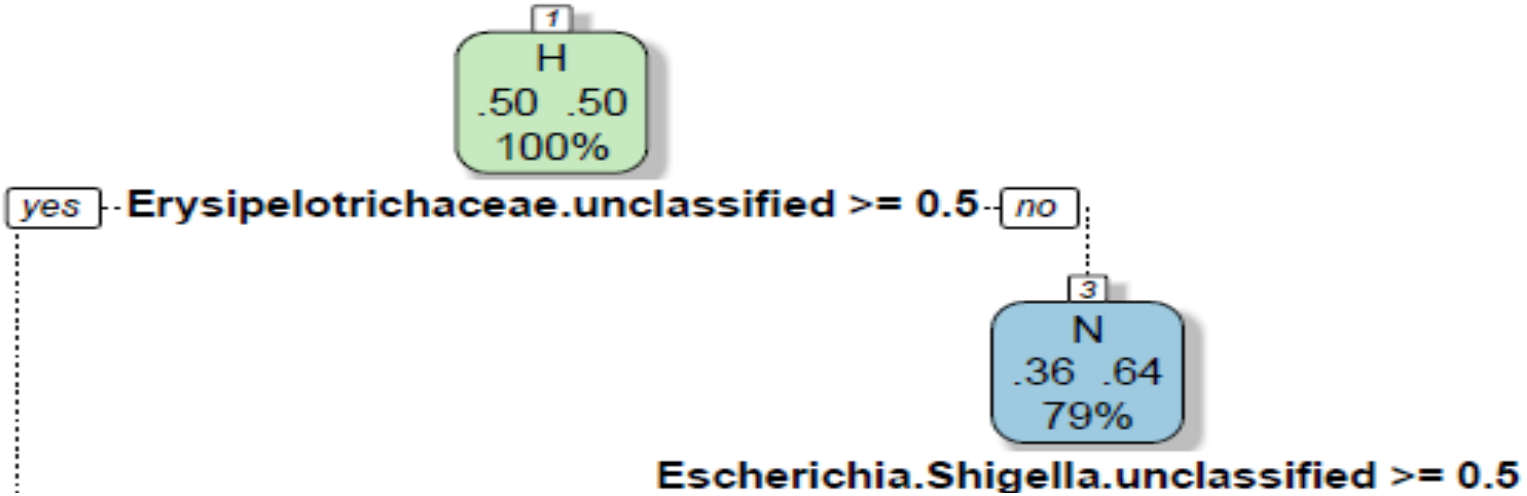
Acibadem Üniversitesi Tıp Fakültesi, Gastroenteroloji Bilim Dalı, Biyoistatistik ve Tıp Bilişimi Ana Bilim Dalı, Tıbbi Mikrobiyoloji Bilim Dalı, İstanbul

Gereç ve Yöntem

- İBS tanı kriterlerini karşılayan, Gastroenteroloji Bilim dalına başvuran 14 hastadan kolonoskopi ile örnek alımı gerçekleşti
- Yaş ve cinsiyet uyumlu tarama amaçlı kolonoskopi yapılan 14 sağlıklı kişiden kontrol grubu oluşturuldu
- Hastalardan ve kontrol gruplarından yaşam tarzı ve yeme alışkanlıkları ile ilgili anket dolduruldu
- 704 taksonomik unit 496 tür elde edildi. Bir grupta diğerine göre 2 kat az veya çok olanların tutulduğu filtre sonrası 30 tür elde edildi. Bu 30 türden bir sınıfta olup diğerinde olmayanlardan ya da az olanlardan bir sınıflama karar ağacı oluşturuldu

E1	E7	E3	E5	F1	F3	F7	Row max	Taxonomic lineages (817 rows)
45	60	15	0	1	14	2	13943	Total reads with no similarities
63108	105754	119864	28444	73025	53242	74027	3799077	Total original input reads
38890	102148	105729	24738	67096	38162	72095	2591870	Total reads mapped to a unique species
4361	0	33943	0	29	0	2015	76665	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterc
1431	69687	265	13110	0	21302	62176	71385	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseud
3523	3846	27	13	53846	986	13	53846	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseud
3279	0	7191	5	3033	1125	79	32327	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__
2980	35	717	477	532	917	136	23270	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Lachno
38	1	21197	0	5	211	0	21197	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__
0	15121	0	1411	410	0	6765	15121	d__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseud
0	1388	0	0	0	0	0	12437	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__
559	8	353	0	114	680	0	12280	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__
926	4	278	0	1	0	0	11201	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__
1842	86	7444	392	133	2536	236	11184	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Rumin
2221	3	364	16	21	83	38	10492	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Lachno
0	0	230	0	0	637	0	9463	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Lachno
239	9	7721	2	137	664	8	7721	d__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__



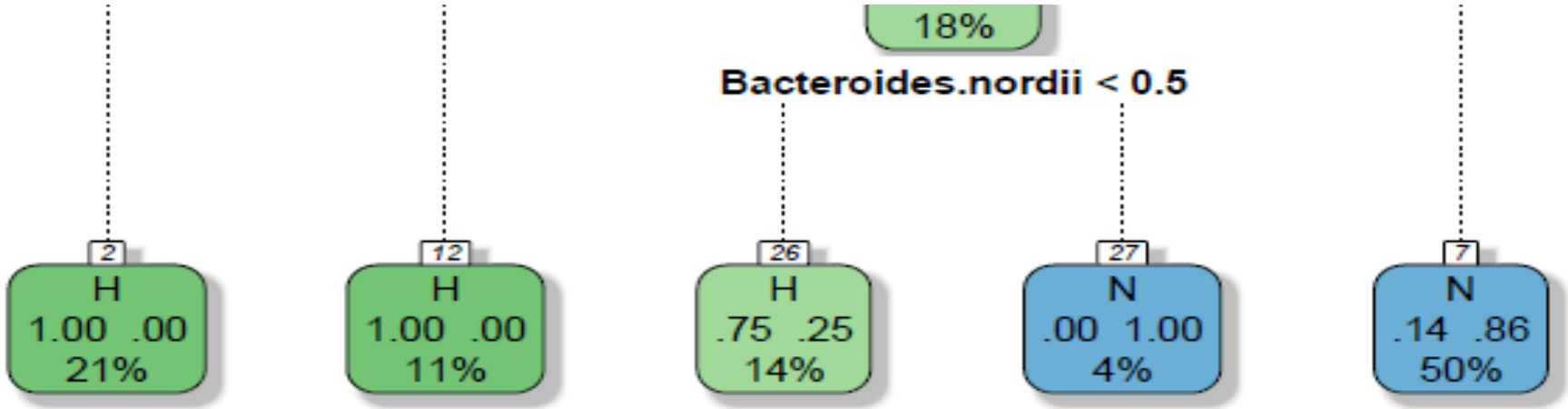


Success! The contingency table below provides the following information: the observed cell totals, (the expected cell totals) and [the chi-square statistic for each cell].

The chi-square statistic, *p*-value and statement of significance appear beneath the table. Blue means you're dealing with dependent variables; red, independent.

	IBS	**Kontrol**	***Marginal Row Totals***
IBS Approved	11 (6.5) [3.12]	2 (6.5) [3.12]	13
Non Detected	3 (7.5) [2.7]	12 (7.5) [2.7]	15
Marginal Column Totals	14	14	28 (Grand Total)

The chi-square statistic is 11.6308. The *p*-value is .000649. This result is significant at $p < .05$.



- THANKS to
- Ege Ülgen
- Burcu Bakır Gungor
- Ozan Ozısık
- Orhan Özcan