University of Zurich[UZH]

BIO392
Bioinformatics of Genome Variations

Genomes: Core of "Personalized Health" & "Precision Medicine"

Michael Baudis **UZH SIB**
Computational Oncogenomics

BIO392: Course Schedule

https://compbiozurich.org/courses/UZH-BIO392/

	Tue Sep 20	Wed Sep 21	Thu Sep 22	Fri Sep 23	Tue Sep 27	Wed Sep 28	Thu Sep 29	Fri Sep 30	Tue Oct 4	Wed Oct 5	Thu Oct 6	Fri Oct 7	Tue Oct 11	Wed Oct 12
09:00 - 10:00			Izaskun: Terminal / Unix / Files	Izaskun: File formats for human genetic variation / file handling		Michael lecture introduction to some resources, CNVs, Progenetix	Hangjia: Blast	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
10:00 - 11:00		Github exercise: create user specific directories & upload/edit test files using Markdown (Ziying)	Izaskun: Terminal / Unix / Files	Izaskun: File formats for human genetic variation / file handling		Task: Browse/explore genome resources and provide some notes (1-2 pages total) in a doc posted on Github (.md)	Hangjia: blast	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
11:00 - 12:00		Ziying: github desktop and terminal		Izaskun: File formats for human genetic variation / file handling			Hangija: Blast exercise	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
13:00 - 14:00	* Room information * Administrative - discuss times/days - exam	Ziying: Introduction to different interfaces eg atom, jupyter, pycharm (lecture), include R things	Izaskun: SIB online introduction to Unix	Izaskun: short project (1000 genomes), reading, literature	Recap W1; Q&A	Hangjia: Progenetix as tool for CNV frequencies etc.	Hangjia: Clinvar and Clingen	Max: STRs	Survival lecture, explanations of terms used etc, cancer classifications	Rahel: Survival analysis	Rahel & Feifei: Survival analysis	Discussion of Survival results (groups of groups)	Exam revision, Q&A	
14:00 - 15:00	Tina Siegenthaler: technical introduction (room, computer, accounts)		Izaskun: SIB online introduction to Unix	Izaskun: short project (1000 genomes), reading, literature	Literature (genome analysis techniques ...)			Max: STRs		Rahel	Rahel & Feifei: Survival analysis			
15:00 - 16:30	* explore course site * create Github accounts and forward to bio392@compbiozurich.org * Michael: short introductory lecture about genome variation			Izaskun: short project (1000 genomes), reading, literature	Genome technologies - brief notes about usage scenarios, pro & con			Max: STRs		Rahel	Rahel & Feifei: Survival analysis			

1992	Heidelberg	Student of medicine | doctoral thesis in molecular cytogenetics @ DKFZ (Peter Lichter) | resident in clinical hematology/oncology | data, clinical studies & cancer systematics
2001	Stanford	Post-doc in hemato-pathology (Michael Cleary) | molecular mechanisms of leukemogenesis | transgenic models | expression arrays | systematic cancer genome data collection | *Progenetix* website
2003	Gainesville	Assistant professor in paediatric haematology | molecular mechanisms of leukemogenesis | focus on bioinformatics for cancer genome data analysis
2006	Aachen	Research group leader in genetics | genomic array analysis for germline alterations | descriptive analysis of copy number aberration patterns in cancer entities
2007	Zürich	Professor of bioinformatics @ DMLS (2015) | systematic assembly of oncogenomic data | databases and software tools | patterns in cancer genomes | *Progenetix* & *arrayMap* resources | GA4GH | SPHN | ELIXIR

Our Research

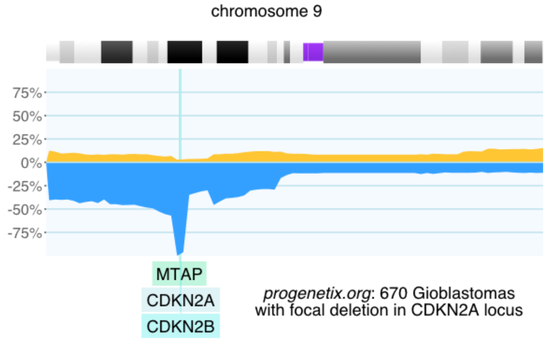
Theoretical Cytogenetics & Oncogenomics

- CNV resource
  - Data - e.g. progenetix.org
  - Tools - CNV remapping, visualization, API access to resources ...
- patterns and correlations of genomic variations in cancer
- annotation mapping
- API, protocols and standards contributions

➡Beacon

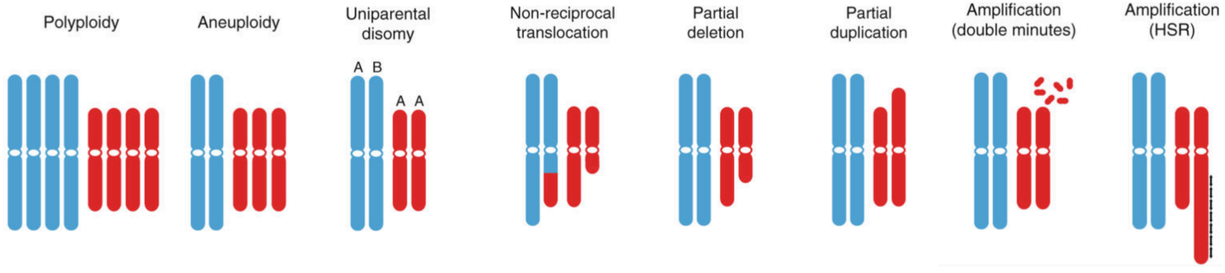
info.baudisgroup.org

Theoretical Cytogenetics and Oncogenomics Research | Methods | Standards

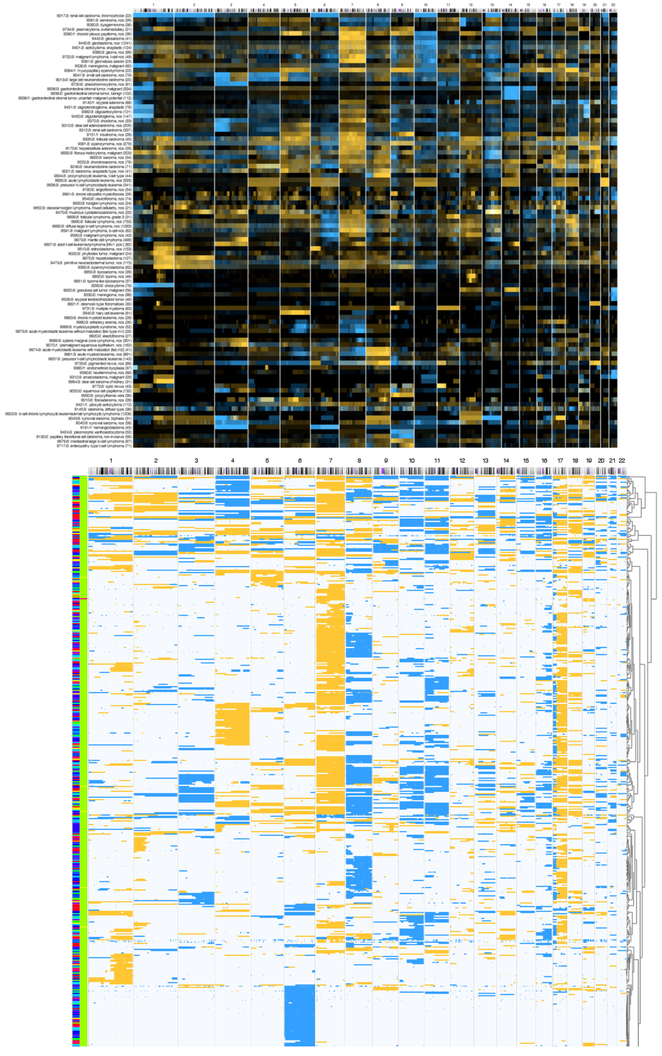


Genomic Imbalances in Cancer - Copy Number Variations (CNV)

- Point mutations (insertions, deletions, substitutions)
- Chromosomal rearrangements
- **Regional Copy Number Alterations** (losses, gains)
- Epigenetic changes (e.g. DNA methylation abnormalities)



Grade et al., 2015 Recent Results Cancer Res



Genome screening at the core of “Personalised Health”

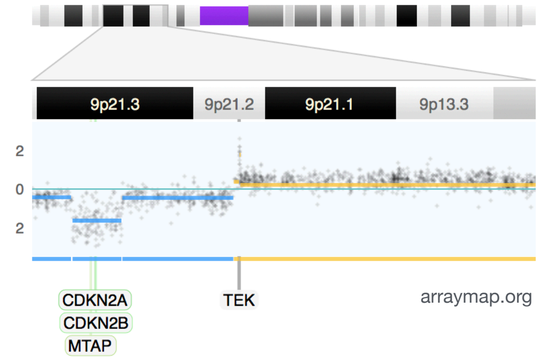
- **Genome analyses** (including transcriptome, metagenomics) are core technologies for Personalised Health™ applications
- The unexpectedly large amount of **sequence variants** in human genomes - germline and somatic/cancer - requires huge analysis efforts and creation of **reference repositories**
- **Standardized data formats** and **exchange protocols** are needed to connect these resources throughout the world, for reciprocal, international **data sharing** and **biocuration** efforts
- Our work @ UZH:
  - ***cancer*** genome repositories
  - biocuration
  - protocols & formats

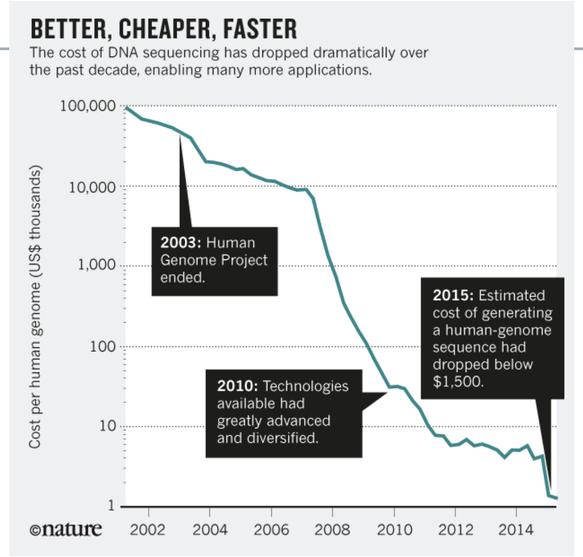
arrayMap

progenetix

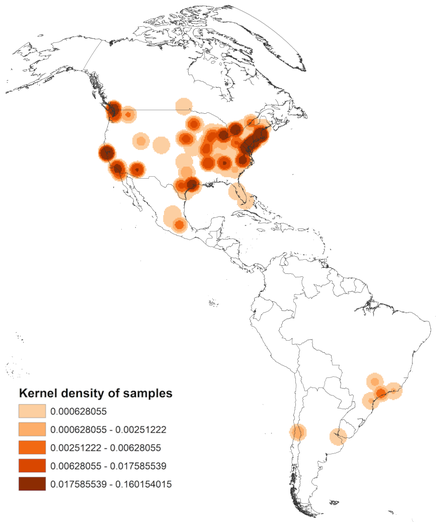
elixir

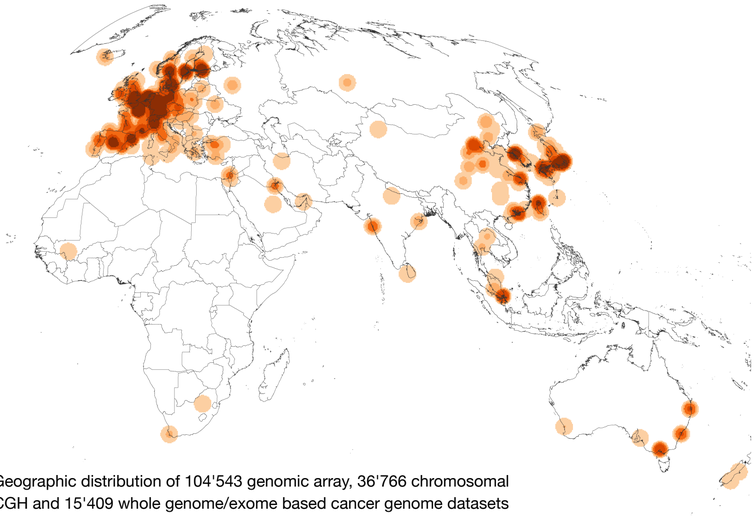
Global Alliance for Genomics & Health





The future of DNA sequencing. Eric D. Green, Edward M. Rubin & Maynard V. Olson. Nature; 11 October 2017 (News & Views)

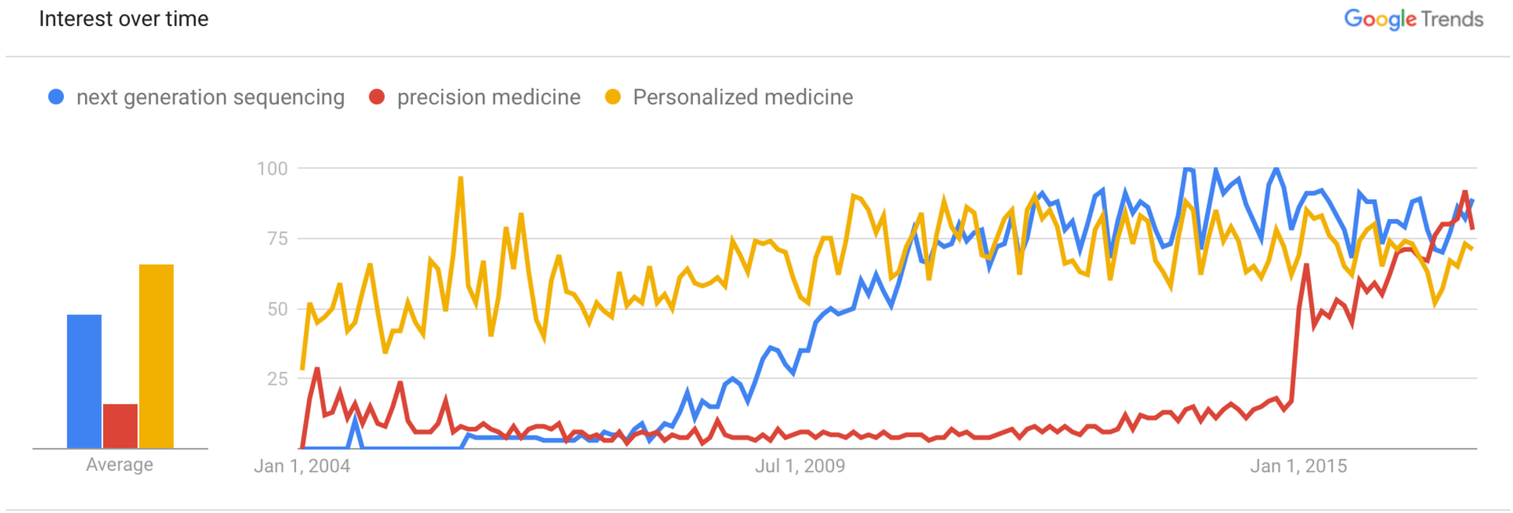




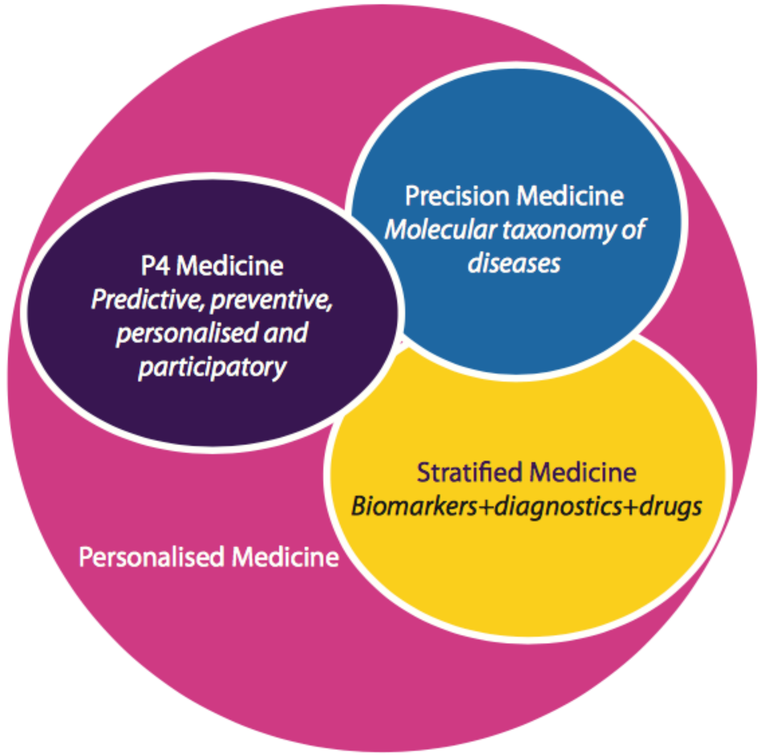
Geographic distribution of 104'543 genomic array, 36'766 chromosomal CGH and 15'409 whole genome/exome based cancer genome datasets

Many names for one concept or many concepts in one name?

Stratified, personalised, precision, individualised, P4 medicine or personalised healthcare – all are terms in use to describe notions often referred to as the future of medicine and healthcare. But what exactly is it all about, and are we all talking about the same thing?



Worldwide. 2004 - present.

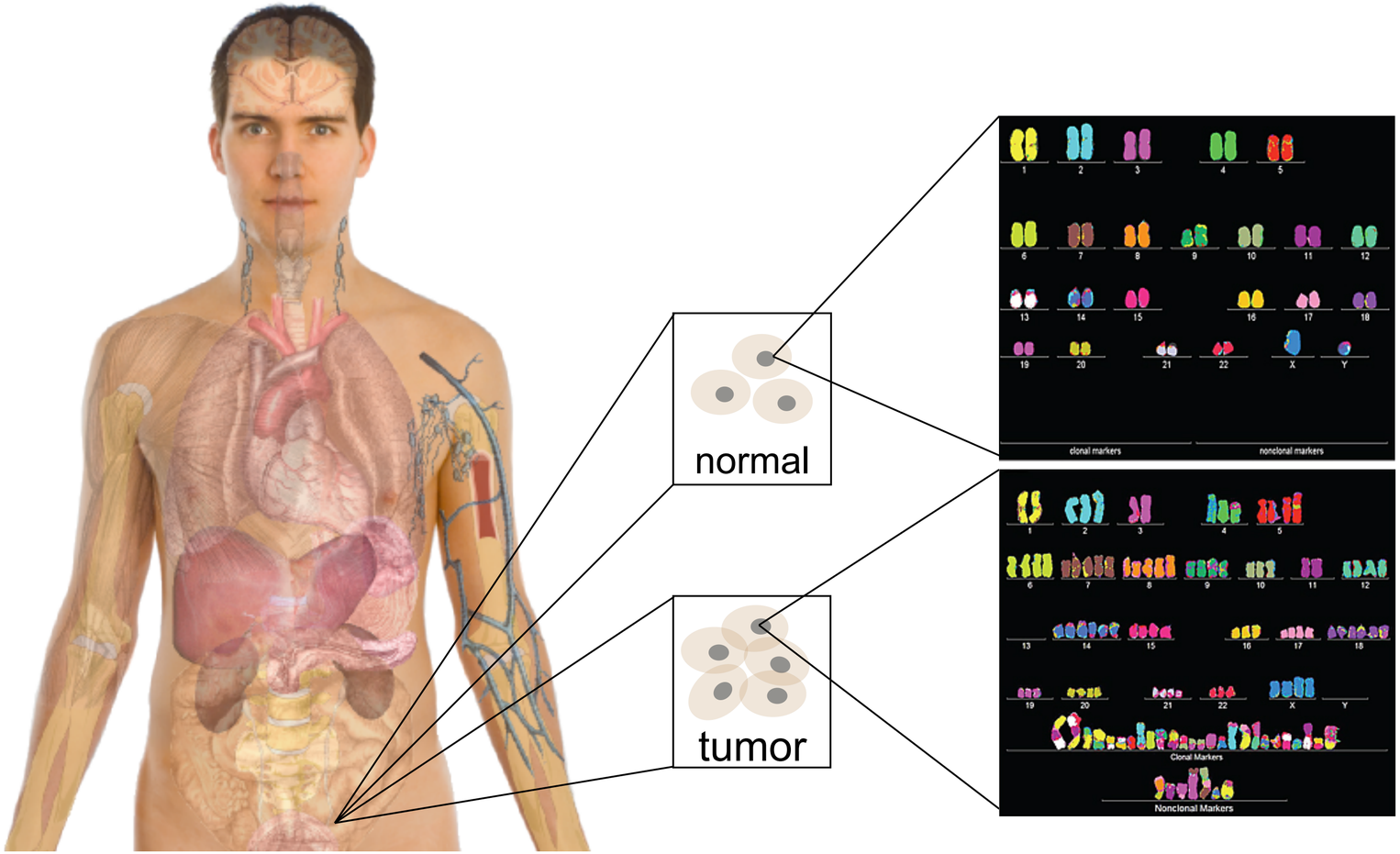


Source: PHG Foundation

While medicine has always been "personal" and "precise" in the given context of available knowledge and technologies, the concept of "**Personalised Medicine**" describes the use of individual genome information, concept based metadata and individually targeted therapies.

Personal Genomics as a Gateway into Biology

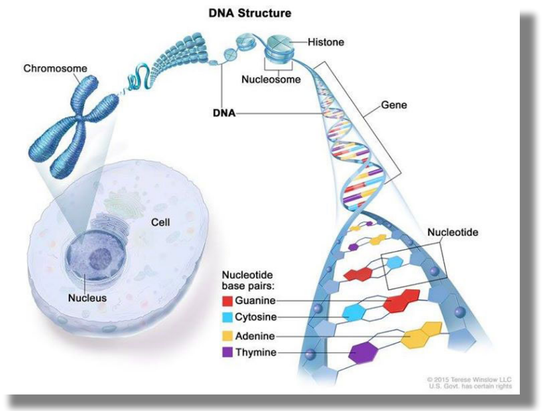
Personal genomes soon will become a commonplace part of medical research & eventually treatment (esp. for cancer). They will provide a primary connection for biological science to the general public.

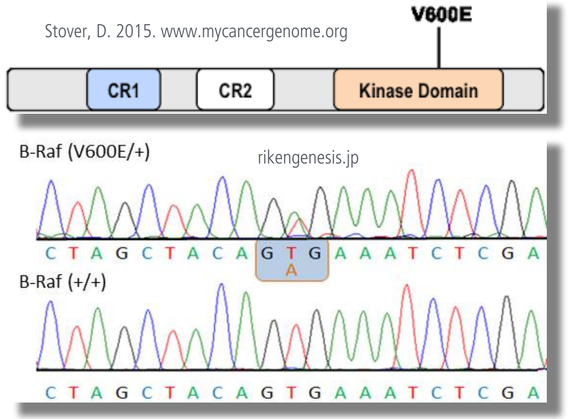


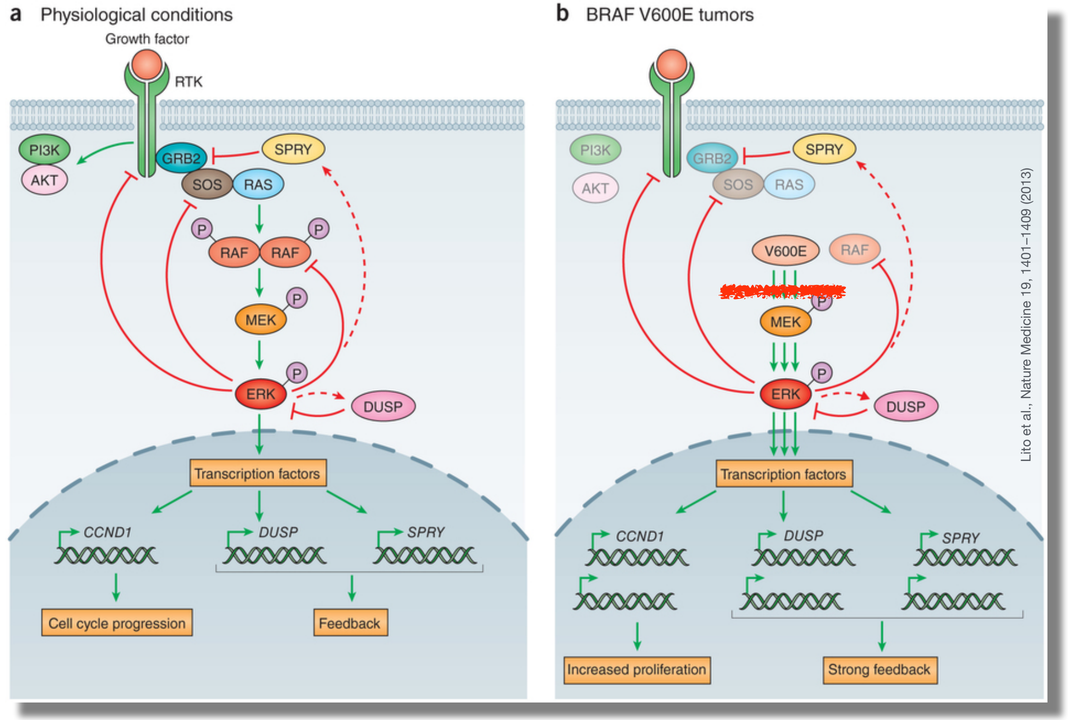
BRAF V600E (c.1799T>A) Mutation

Oncogene Activation by Single Nucleotide Alteration

- a single nucleotide exchange **T**hymidine > **A**denine leads to continuous RAF based activation of the MEK-ERK pathway
- BRAF V600E is a frequent mutation in >50% of malignant melanomas, but also CRC, lung ADC …
- pharmacologic block of B-Raf (e.g. through ***Vemurafenib***)







The BRAF V600E mutation leads to continuos phosphorylation of MEK, without the need for receptor based activation of the upstream pathway and loss of inhibitory feedback control.

Genome analyses at the core of Personalized Health™

Susceptibility, Pharmacogenomics, Classification, Infectious Diseases, Outcome Prediction, Lifestyle ...

doi:10.1038/nature19057

Analysis of protein-coding genetic variation in 60,706 humans

Monkol Lek[1,2,3,4], Konrad J. Karczewski[1,2]*, Eric V. Minikel[1,2,5]*, Kaitlin E. Samocha[1,2,5,6]*, Eric Banks[2], Timothy Fennell[2], Anne H. O'Donnell-Luria[1,2,7], James S. Ware[2,8,9,10,11], Andrew J. Hill[1,2,12], Beryl B. Cummings[1,2,5], Taru Tukiainen[1,2], Daniel P. Birnbaum[2], Jack A. Kosmicki[1,2,6,13], Laramie E. Duncan[1,2,6], Karol Estra

Barkur S. Shastry

SNP alleles in human disease and evolution

DISEASE MECHANISMS

Mechanisms underlying structural variant formation in genomic disorders

Claudia M. B. Carvalho[1,2] and James R. Lupski[1,3,4,5]

Abstract | With the recent burst of technological developments in genomics, and the clinical implementation of genome-wide assays, our understanding of the molecular basis of genomic disorders, specifically the contribution of structural variation to disease burden, is evolving

Rapid whole genome sequencing and precision neonatology

Joshua E. Petrikin, MD[a,], Laurel K. Willig, MD, FAAP[b], Laurie D. Smith, MD, PhD[c], and Stephen F. Kingsmore, MB, BAO, ChB, Dsc, FRCPath[d,e]*

insight progress

Cancer genetics

Bruce A. J. Ponder

Genomic Classification of Cutaneous Melanoma

The Cancer Genome Atlas Network[1,*,*,*]
[1]Cancer Genome Atlas Program Office, National Cancer Institute at NIH, 31 Center Drive, Bldg. 31, Suite 3A20, Bethesda, MD 20892, USA
*Correspondence: irwatson@mdanderson.org (I.R.W.), jgershen@mdanderson.org (J.E.G.), lchin@mdanderson.org (L.C.)
http://dx.doi.org/10.1016/j.cell.2015.05.044

Consequences of genomic diversity in *Mycobacterium tuberculosis*

Mireia Coscolla[a,b], Sebastien Gagneux[a,b,*]

[a] *Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4002 Basel, Switzerland*
[b] *University of Basel, Petersplatz 1, Basel 4003, Switzerland*

doi:10.1111/pcn.12128

Psychiatry and Clinical Neurosciences

PCN Frontier Review

Copy-number variation in the pathogenesis of autism spectrum disorder

Emiko Shishido, PhD[1,2,3], Branko Aleksic, MD, PhD[3] and Norio Ozaki, MD, PhD[3]*

he Promotion of Science, Japan

Common gene variants, mortality and extreme longevity in humans

B.T. Heijmans[a,b], R.G.J. Westendorp[b], P.E. Slagboom[a,*]

RESEARCH ARTICLE Open Access

Integrative genome-wide expression profiling identifies three distinct molecular subgroups of renal cell carcinoma with different patient utcome

nfred Beleut[1,5*], Philip Zimmermann[2], Michael Baudis[3], Nicole Bruni[4], Peter Bühlmann[4], Oliver Laule[2], -Duc Luu[1], Wilhelm Gruissem[2], Peter Schraml[1*] and Holger Moch[1]

NEURODEVELOPMENT

Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders

Mustafa Sahin* and Mriganka Sur*

The landscape of somatic copy-number alteration across human cancers

Rameen Beroukhim[1,3,4,5]*, Craig H. Mermel[1,3]*, Dale Porter[8], Guo Wei[1], Soumya Raychaudhuri[1,4], Jerry Donovan[8], Jordi Barretina[1,3], Jesse S. Boehm[1], Jennifer Dobson[1,3], Mitsuyoshi Urashima[9], Kevin T. Mc Henry[8], Reid M. Pinchback[1], Azra H. Ligon[4], Yoon-Jae Cho[6], Leila Haery[1,3], Heidi Greulich[1,3,4,5], Michael Reich[1], Wendy Winckler[1], Michael S. Lawrence[1], Barbara A. Weir[1,3], Kumiko E. Tanaka[1,3], Derek Y. Chiang[1,3,13], Adam J. Bass[1,3,4], Alice Loo[8], Carter Hoffman[1,3], John Prensner[1,3], Ted Liefeld[1], Qing Gao[1], Derek Yecies[3], Sabina Signoretti[3,4], Elizabeth Maher[10], Frederic J. Kaye[11], Hidefumi Sasaki[12], Joel E. Tepper[13], Jonathan A. Fletcher[4], Josep Tabernero[14], José Baselga[14], Ming-Sound Tsao[15], Francesca Demichelis[16], Mark A. Rubin[16], Pasi A. Janne[3,4], Mark J. Daly[1,17], Carmelo Nucera[7], Ross L. Levine[18], Benjamin L. Ebert[1,4,5], Stacey Gabriel[1], Anil K. Rustgi[19], Cristina R. Antonescu[18], Marc Ladanyi[18], Anthony Letai[3], Levi A. Garraway[1,3], Massimo Loda[3,4], David G. Beer[20], Lawrence D. True[21], Aikou Okamoto[22], Scott L. Pomeroy[6], Samuel Singer[18], Todd R. Golub[1,3,23], Eric S. Lander[1,2,5], Gad Getz[1], William R. Sellers[8] & Matthew Meyerson[1,3,5]

RESEARCH ARTICLE Open Access

Chromothripsis-like patterns are recurring but heterogeneously distributed features in a survey of 22,347 cancer genome screens

Haoyang Cai[1,2], Nitin Kumar[1,2], Homayoun C Bagheri[3], Christian von Mering[1,2], Mark D Robinson[1,2]* and Michael Baudis[1,2]*

Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib

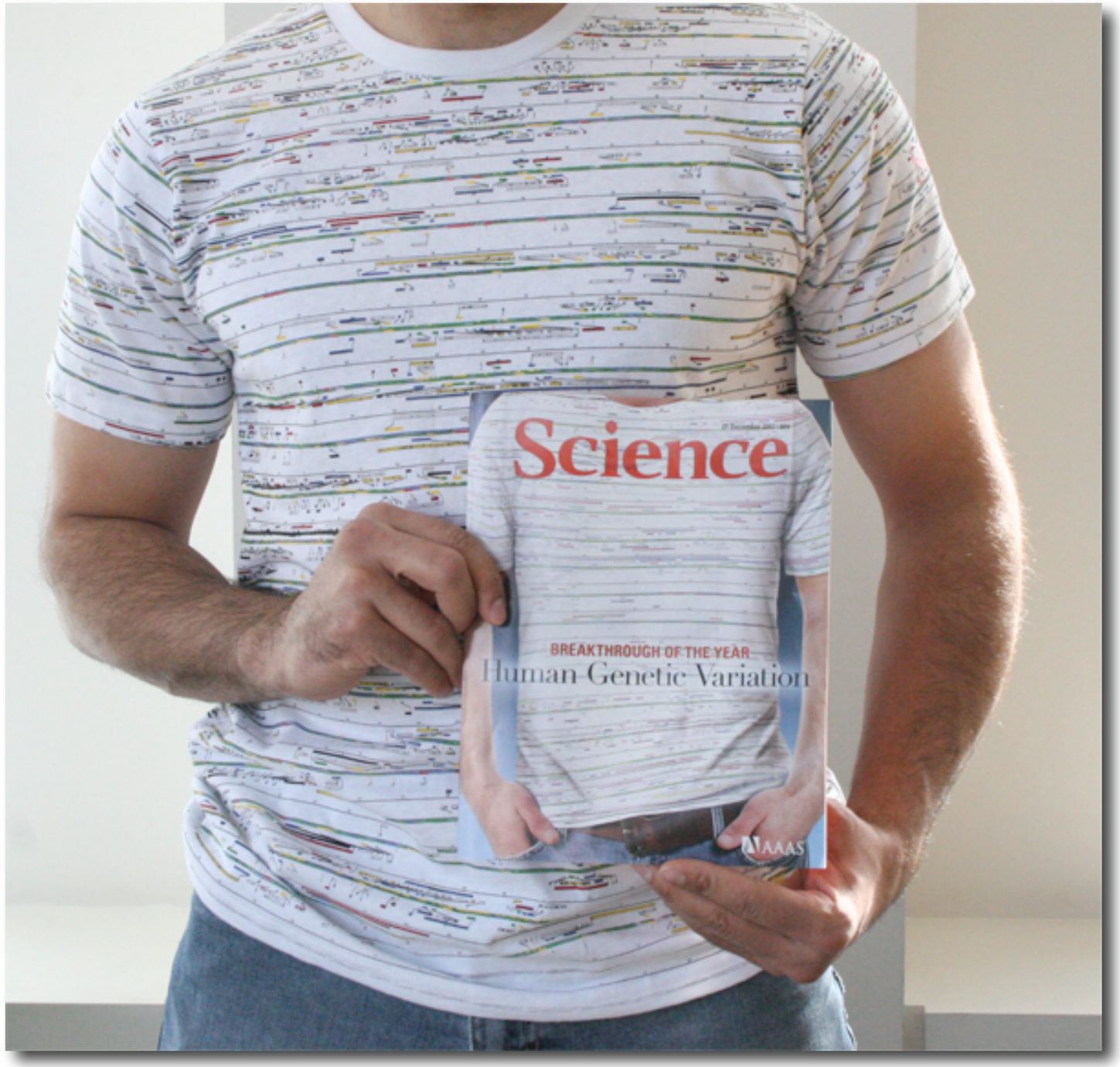
Thomas J. Lynch, M.D., Daphne W. Bell, Ph.D., Raffaella Sordella, Ph.D., Sarada Gurubhagavatula, M.D., Ross A. Okimoto, B.S., Brian W. Brannigan, B.A., Patricia L. Harris, M.S., Sara M. Haserlat, B.A., Jeffrey G. Supko, Ph.D., Frank G. Haluska, M.D., Ph.D., David N. Louis, M.D., David C. Christiani, M.D., Jeff Settleman, Ph.D., and Daniel A. Haber, M.D., Ph.D.
N Engl J Med 2004; 350:2129-2139 | May 20, 2004 | DOI: 10.1056/NEJMoa040938

Genome analyses at the core of Personalized Health™

There'll be Sequencing Everywhere...

- Genome analyses (including transcriptome, metagenomics) are the **core technologies** for Personalized Health™ applications
- In the context of **academic medicine**, this requires
  - standard sample acquisition procedures & central **biobanking**
  - **core sequencing facility** (large throughput, cost efficiency, uniform sample and data handling procedures)
- secure **computing/analysis** platform
- Standardized **data formats** and **sample identification** procedures
- Metadata rich, reference **variant resource(s)** & expertise
- participation in reciprocal, international **data sharing** and **biocuration** efforts

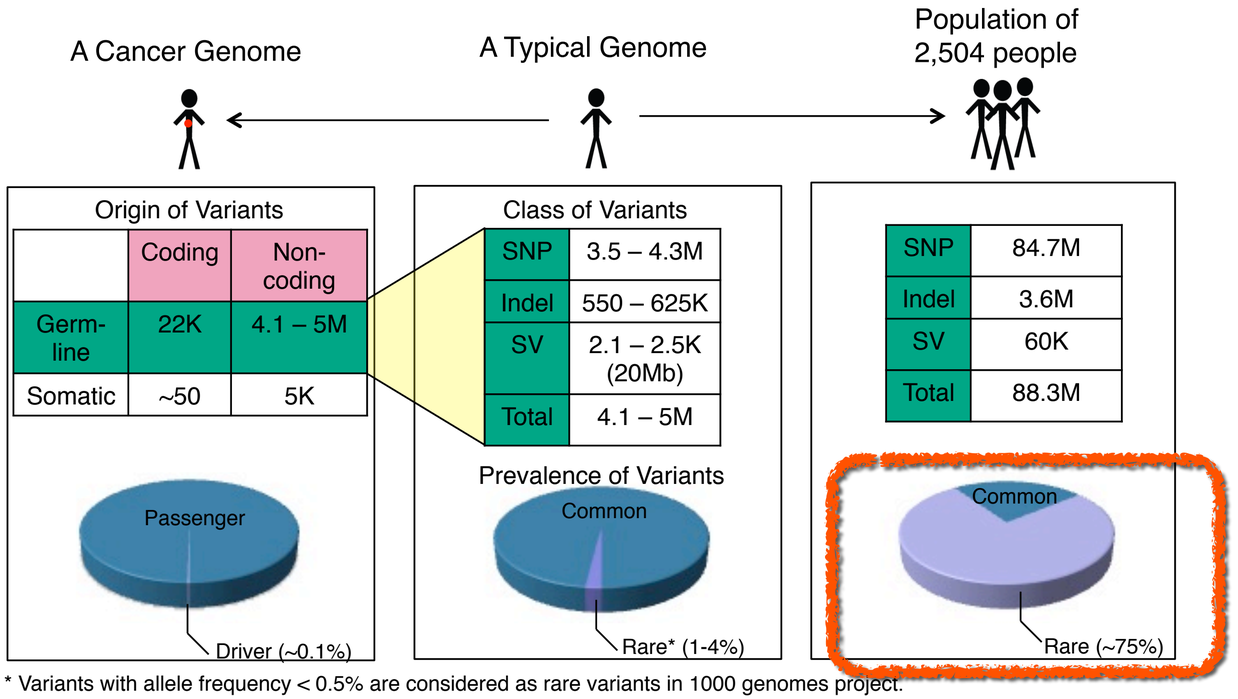
The trouble with human genome variation



Finding Somatic Mutations In Cancer

Many Needles in a Large Haystack

- a typical human genome (~3 billion base pairs) has ~5 million variants
- most of them are "**rare**"; i.e. can only be identified as recurring when sequencing thousands of people
- cancer cells accumulate additional variants, only **few** of which ("**drivers**") are relevant for the disease

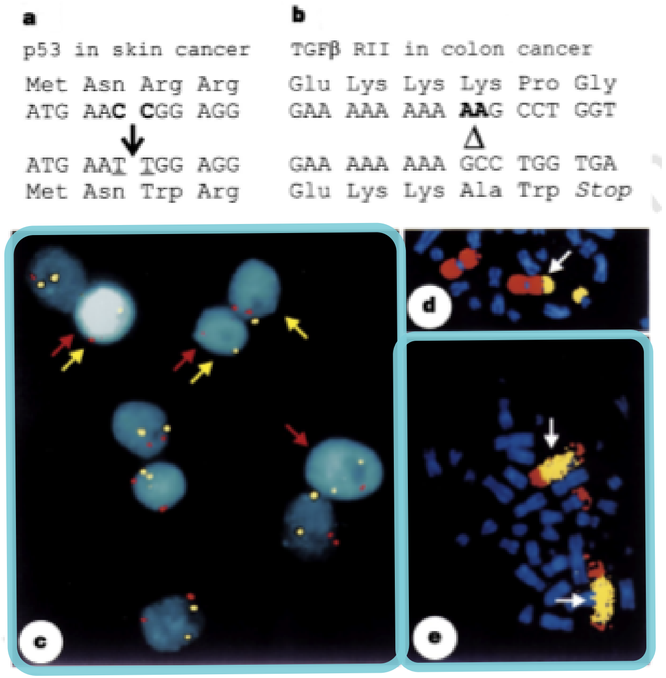


* Variants with allele frequency < 0.5% are considered as rare variants in 1000 genomes project.

The 1000 Genomes Project Consortium, Nature. 2015. 526:68-74
Khurana E. et al. Nat. Rev. Genet. 2016. 17:93-108

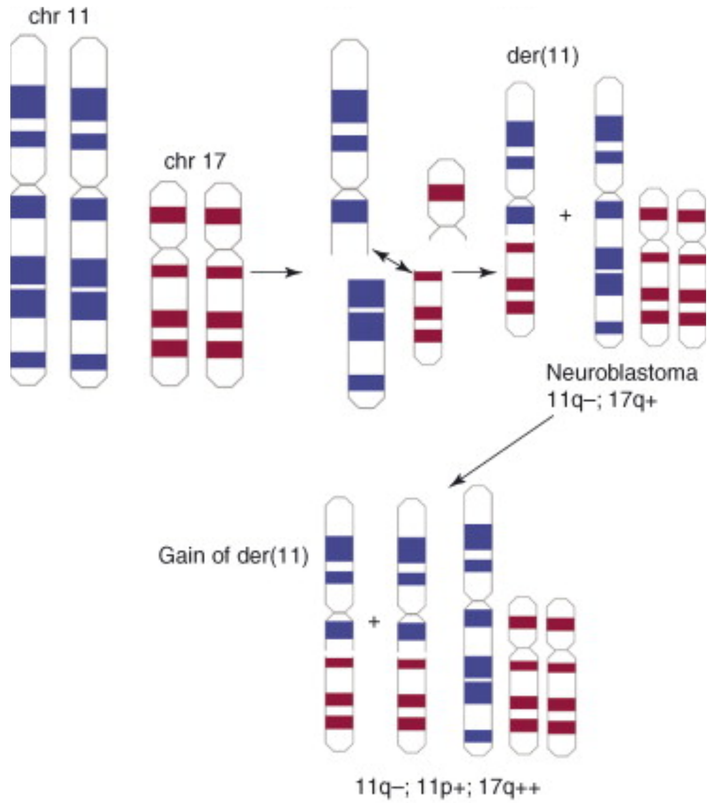
Graphic adapted from Mark Gerstein (GersteinLab.org; @markgerstein)

Mutations & genomic rearrangements in cancer



Lengauer et al. Genetic instabilities in human cancers. Nature (1998) vol. 396 (6712) pp. 643-9

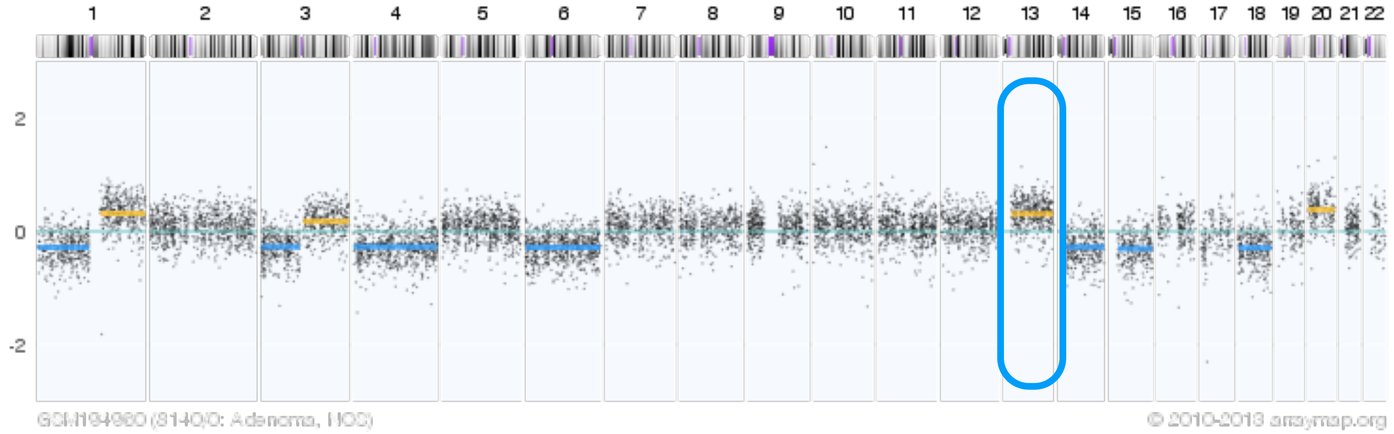
a. small mutation (di-pyrimidine exchange at p53 in Xeroderma pigmentosum patient)
b. two-base deletion in *TGFB* in a colorectal cancer patient with mismatch repair deficiency
c. chromosomal losses (FISH; red=3, yellow=12) in CRC
d. t(1;17) in neuroblastoma, whole-chromosomal painting
e. *MYCN* gene amplification (multiple copies inserted into chromosome 1 derived marker)



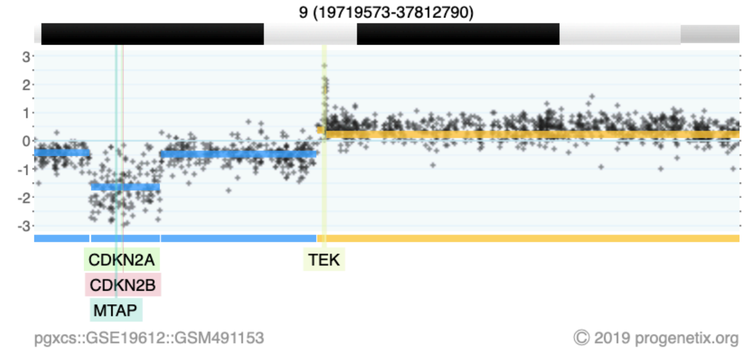
Generation of copy number imbalances in cancer through imbalanced cytogenetic rearrangements - partial deletion of 11q, gain of 11pterq21 and 2 addl. copies of 17q

RL Stallings: Are chromosomal imbalances important in cancer? Volume 23, Issue 6, p278–283, 2007

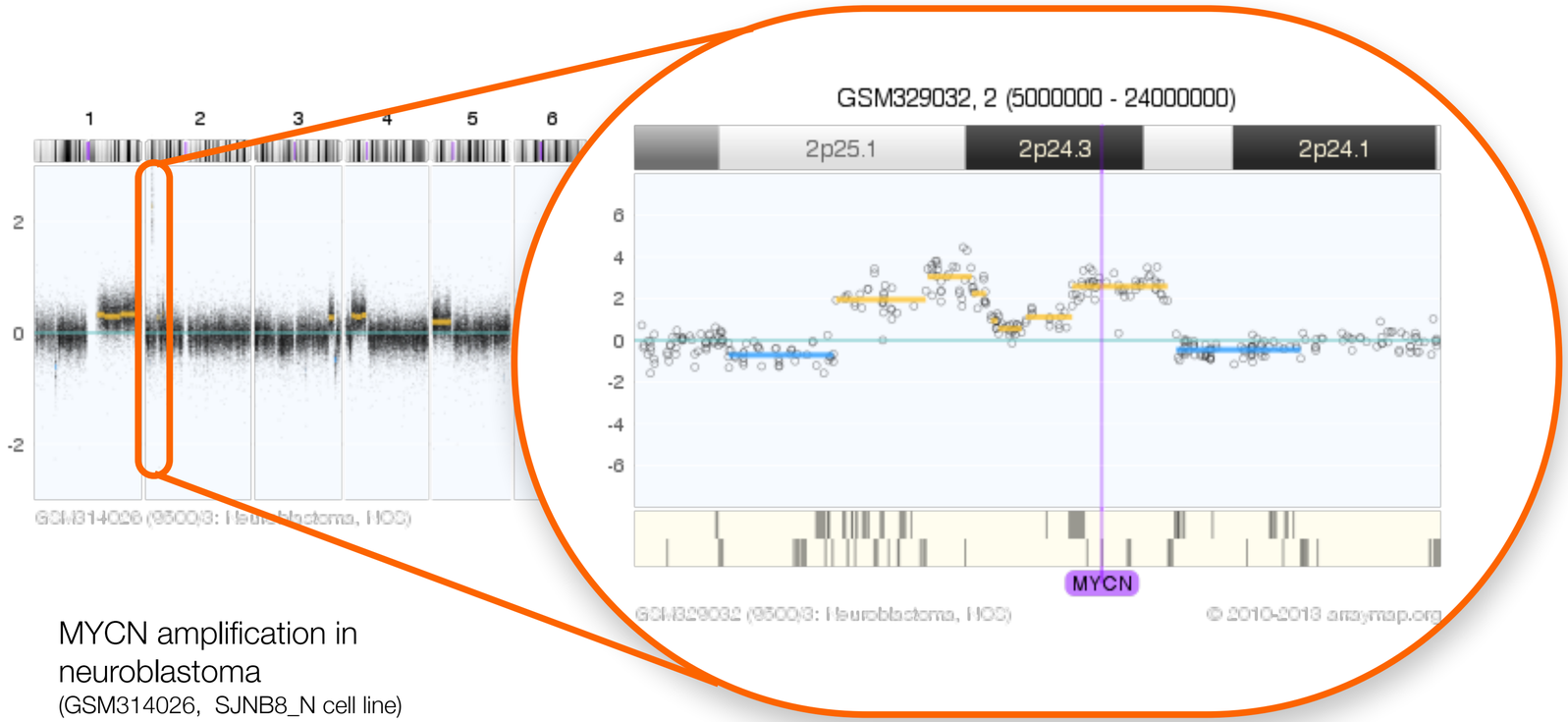
Somatic Copy Number Variations



Gain of chromosome arm 13q in colorectal carcinoma



2-event, homozygous deletion in a Glioblastoma

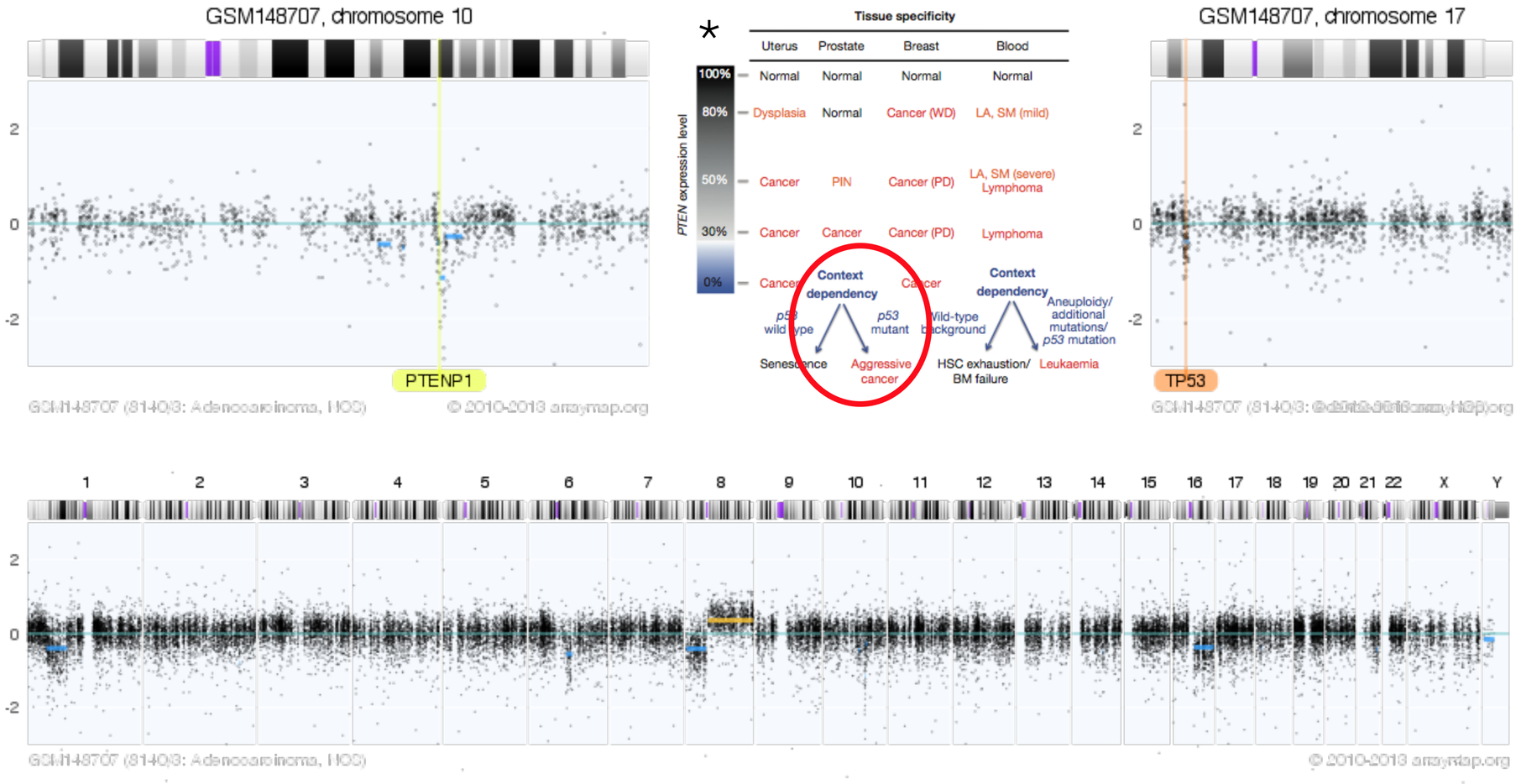


MYCN amplification in neuroblastoma
(GSM314026, SJNB8_N cell line)

low level/**high level** copy number alterations (CNAs)

arrayMap

Gene dosage phenomena beyond simple on/off effects

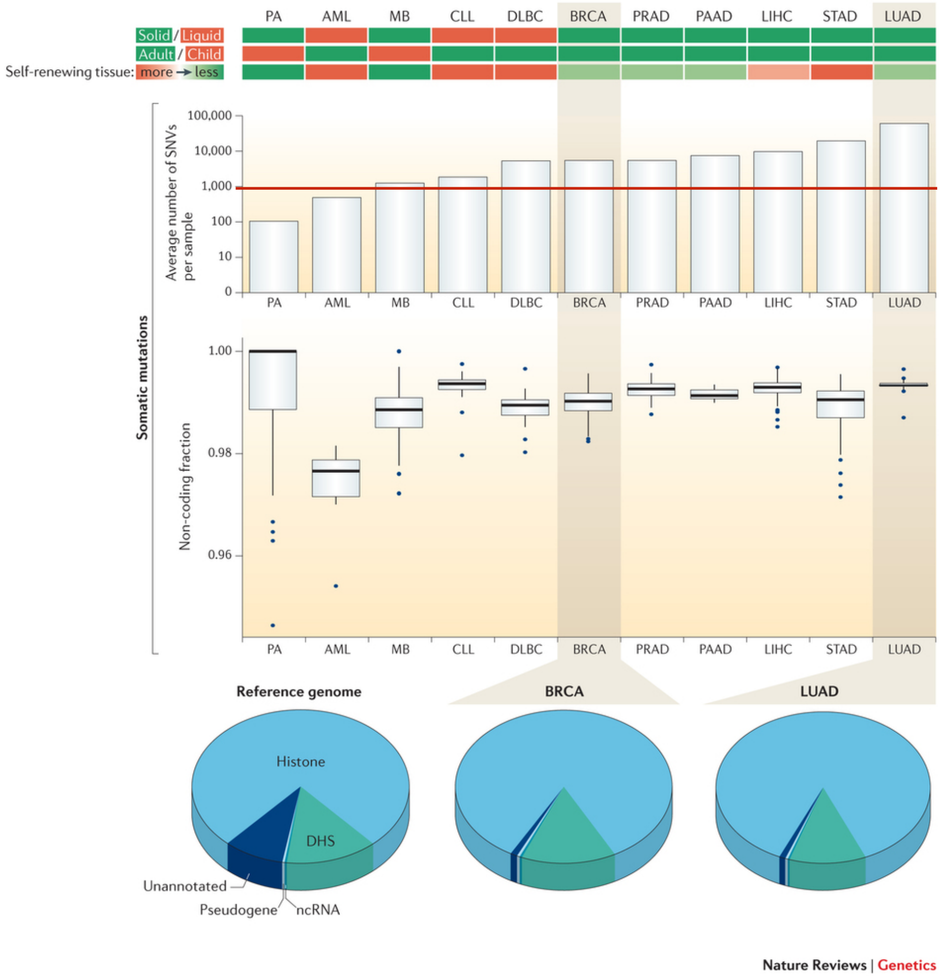


Combined heterozygous deletions involving *PTEN* and TP53 loci in a case of prostate adenocarcinoma
(GSM148707, PMID 17875689, Lapointe *et al*., CancRes 2007)

* A. H. Berger, A. G. Knudson, and P. P. Pandolfi, “A continuum model for tumour suppression,” *Nature*, vol. 476, no. 7359, pp. 163–169, Aug. 2011.

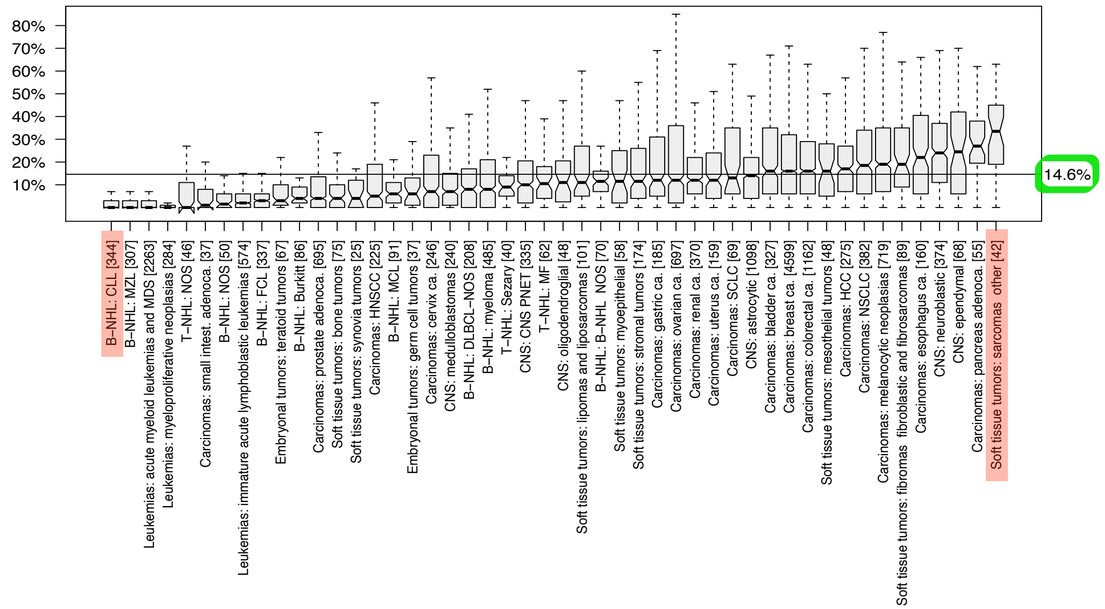
arrayMap

Quantifying Somatic Mutations In Cancer



CANCERS SHOW THOUSANDS OF SINGLE NUCLEOTIDE VARIANTS PER SAMPLE, MOSTLY IN NON-CODING REGIONS

Pan-Cancer Analysis of Whole Genomes (PCAWG) data show widespread mutations in non-coding regions of cancer genomes (Khurana et al., Nat. Rev. Genet. (2016)

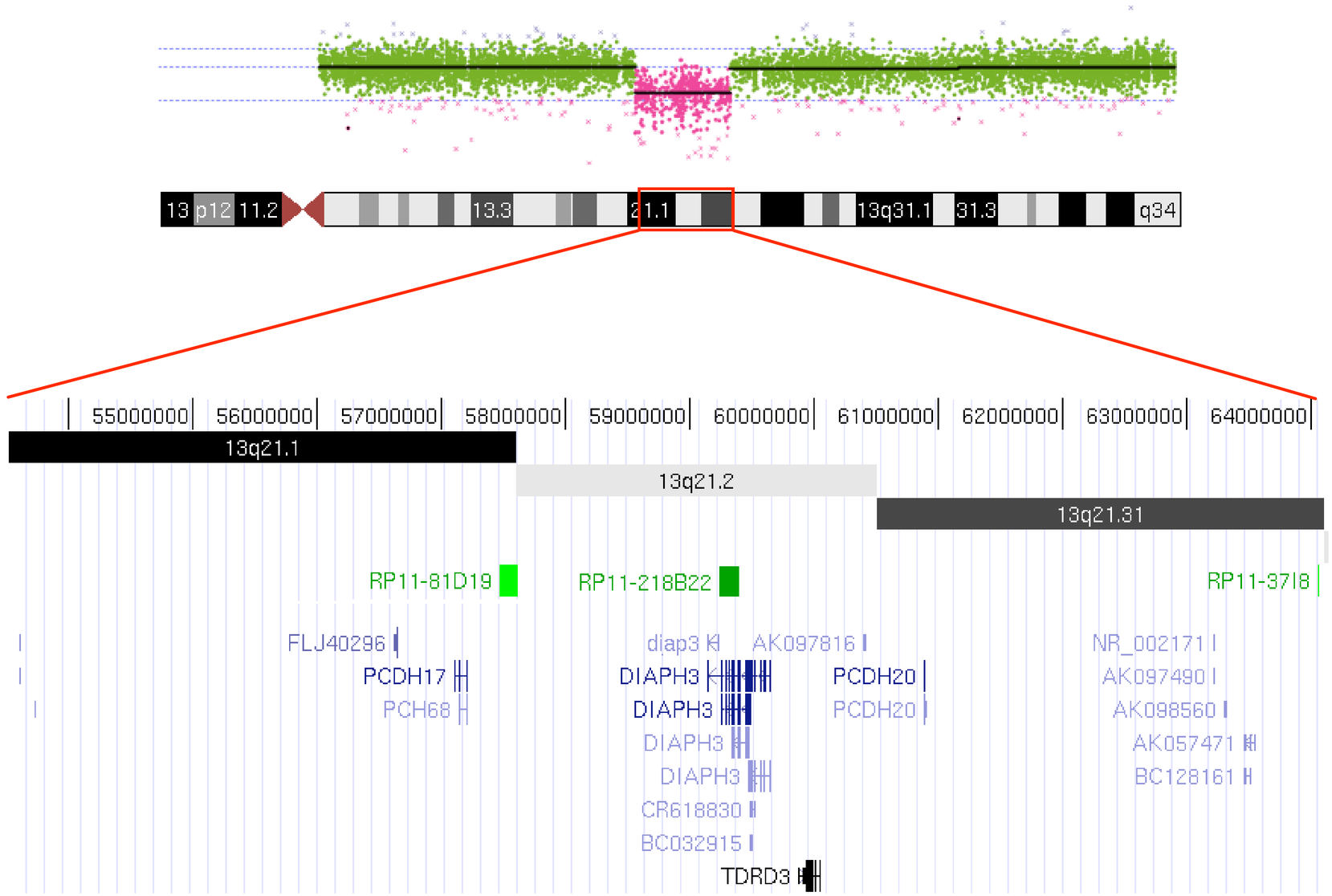


GENOMIC COPY NUMBER IMBALANCES PROVIDE WIDESPREAD SOMATIC VARIANTS IN CANCER

On average ~15% of a cancer genome are in an imbalanced state (more/less than 2 alleles); Original data based on >30'000 cancer genomes from arraymap.org

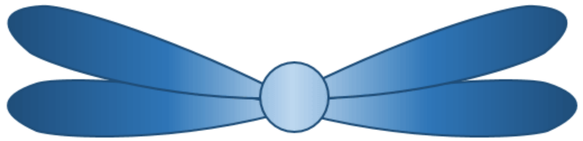
Nobody is perfect (?)

A 10.7 Mb Interstitial Deletion of 13q21 Without Phenotypic Effect Defines a Further Non-Pathogenic Euchromatic Variant
Andreas Roos, Miriam Elbracht, Michael Baudis, Jan Senderek, Nadine Schönherr, Thomas Eggermann, and Herdit M. Schüler
American Journal of Medical Genetics Part A 146A:2417 – 2420 (2008)

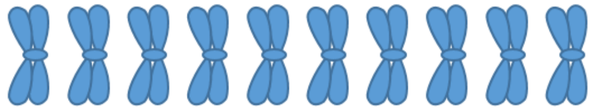


Genome Sequencing

whole genome sequencing (WGS)

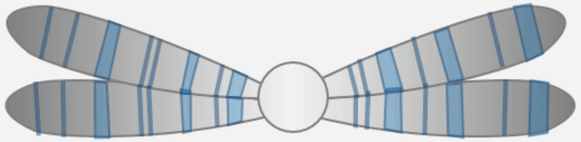


100%



all DNA
(3.1 billion base pairs)

exome sequencing



~1%

protein-coding DNA only
(~31 million base pairs)

What does it cost to sequence a genome?

	Human Genome Project (HGP): 1991-2003	today: 2017
cost:	$2.7 billion	~$1,500
time:	12+ years	< 2 days

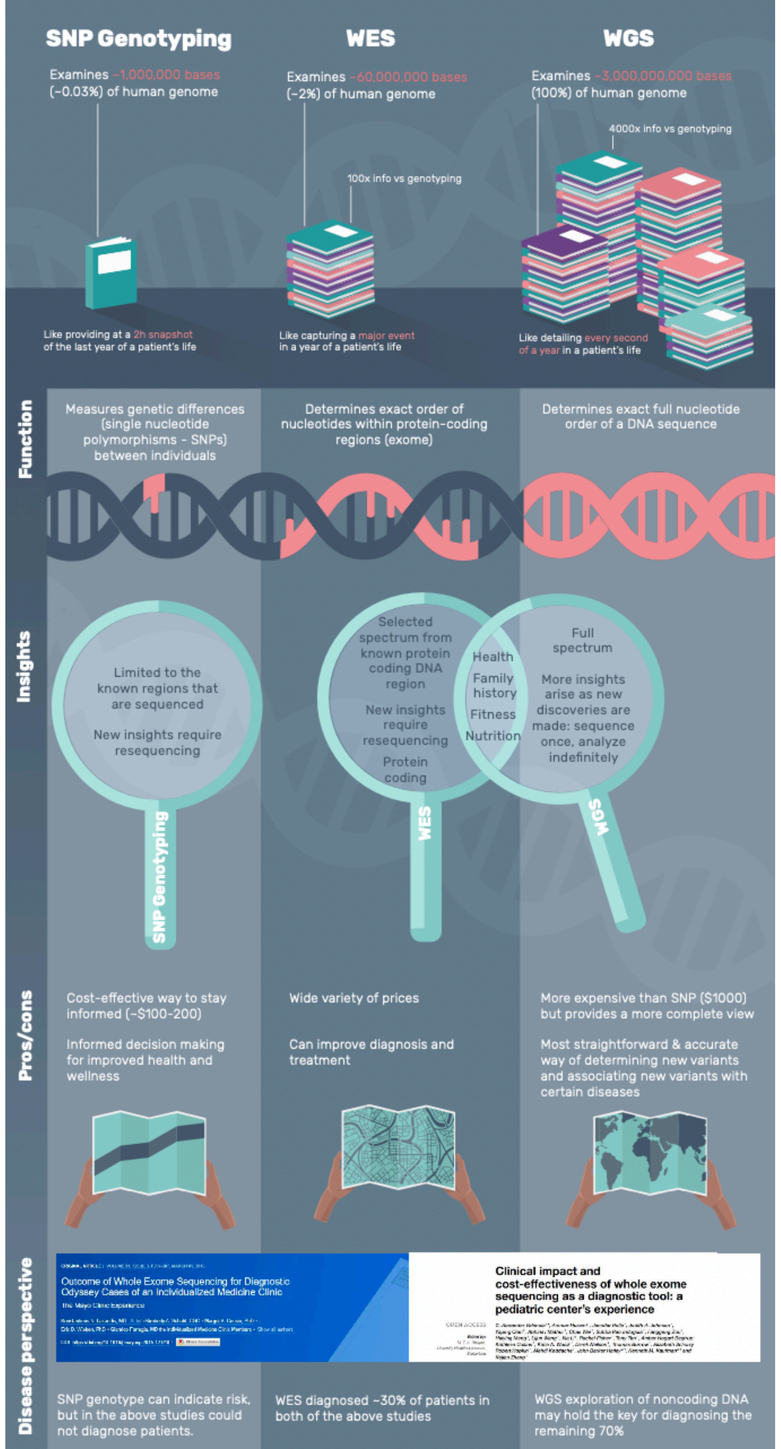
today:
2017

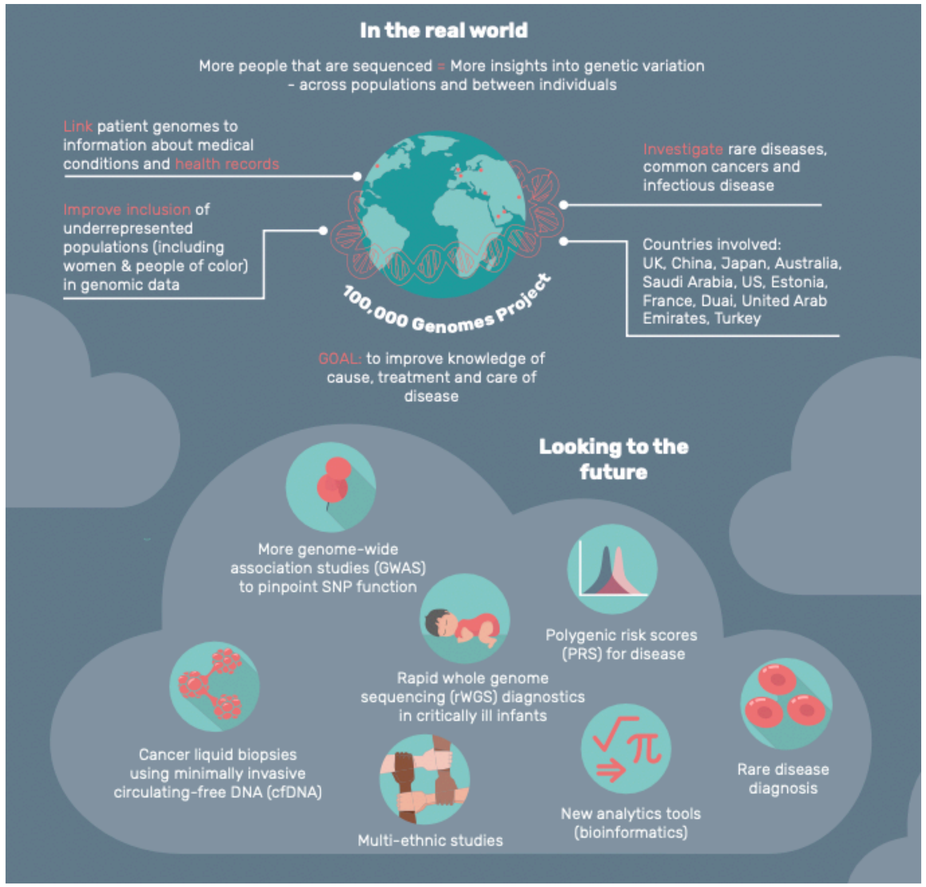
~$530
~3 days

Ryan L. Collins & Brad Wierbowski (2017)

Genome Analysis

A “progressing technologies” view

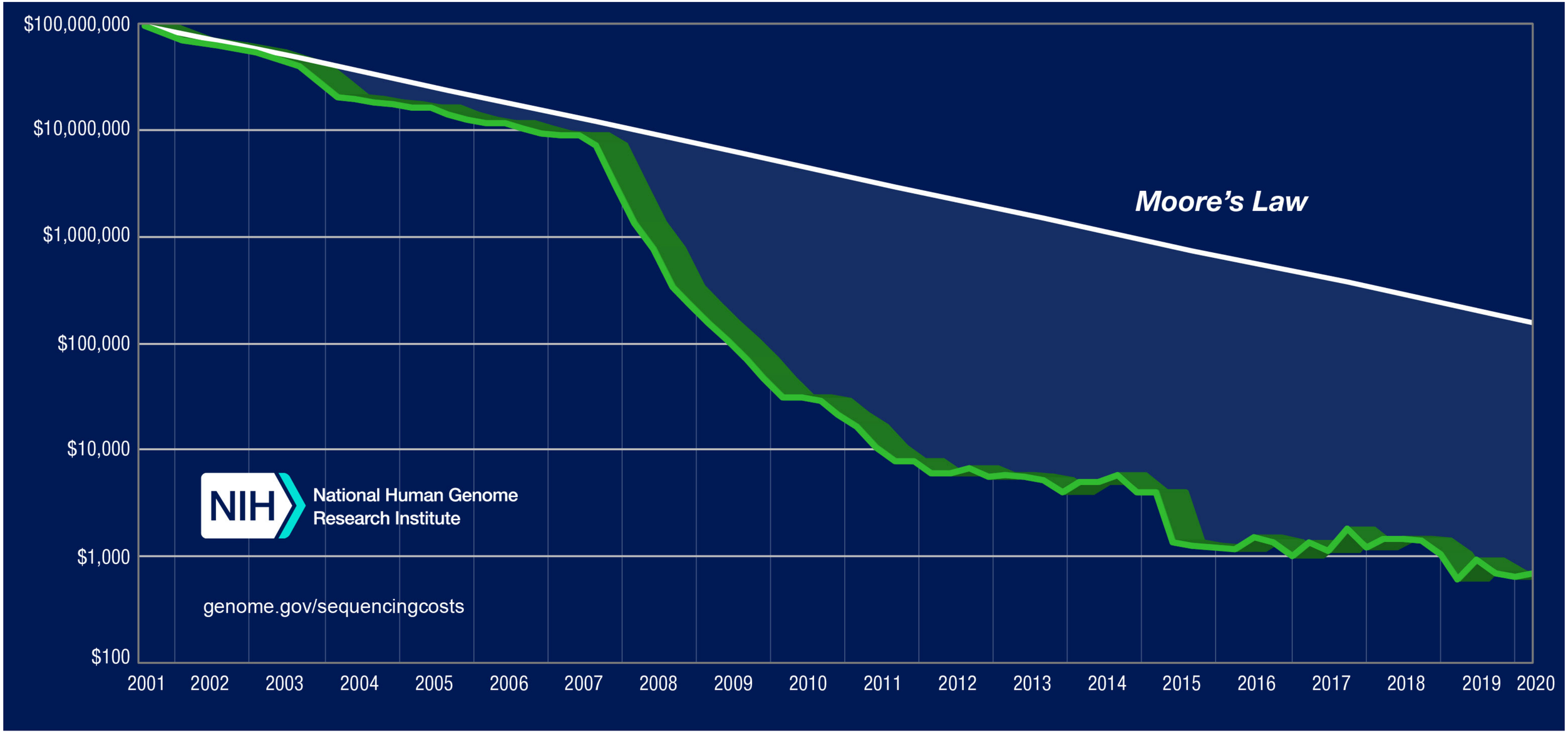




BioTechniques

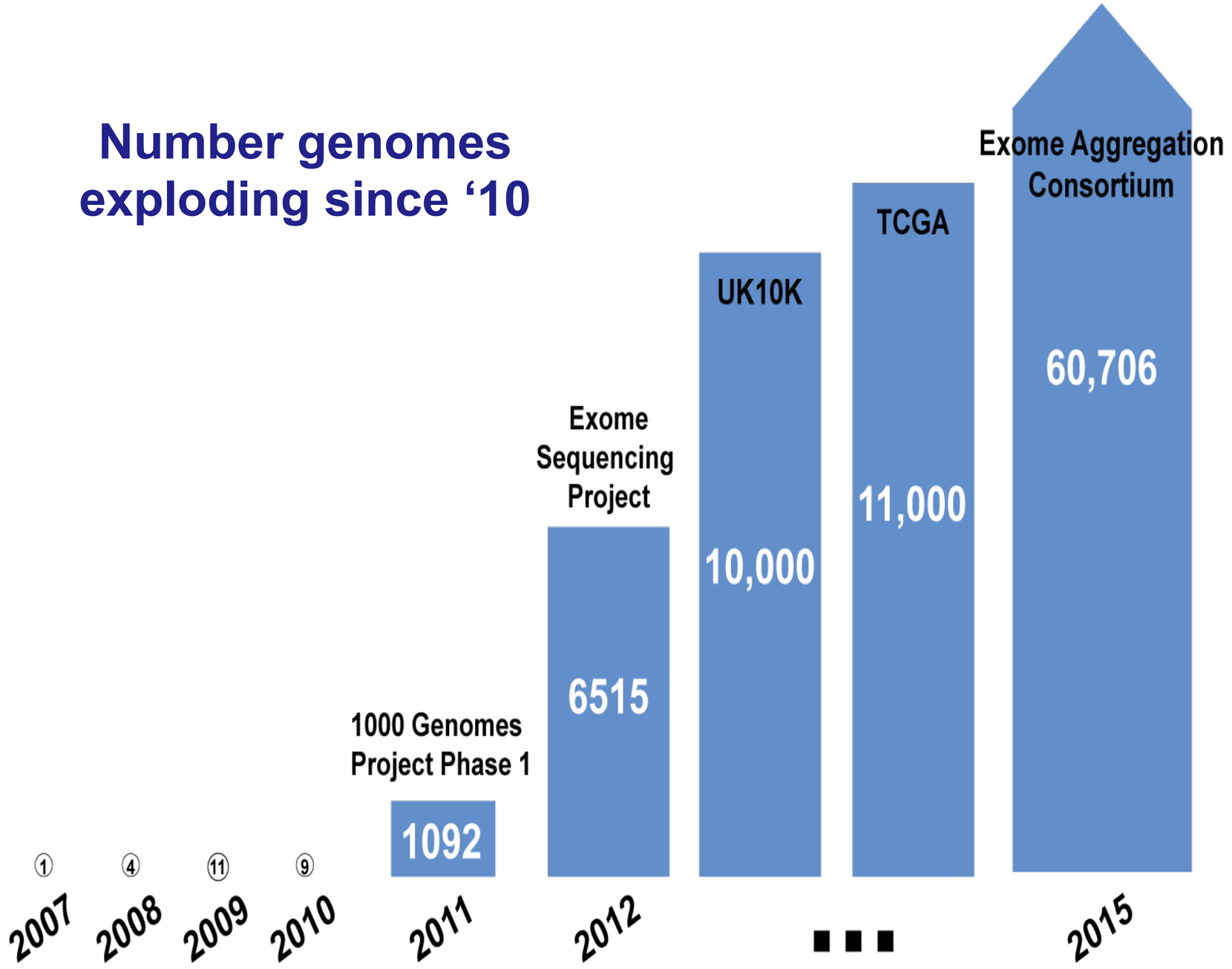
psomagen

The Cost of Sequencing a Human Genome



Slides freely downloadable from lectures.GersteinLab.org & “tweetable” (via @markgerstein)

Number genomes exploding since ‘10



Genomes Everywhere

Large Genome Data Generation, Analysis & Sharing Initiatives

Organization / Initiative: Name	Organization / Initiative: Category	Cohort
100K Wellness Project	Research Project	107 unaffected individuals (scaling up to 100,000)
23andMe	Organization	>1 million customers (>80% consented to research)
Actionable Cancer Genome Initiative (ACGI)	Data-Sharing Project	Goal: 100,000 individuals
Ancestry.com	Organization	1.4 million customer DNA samples (what % consented to research?)
BioBank Japan	Repository	Specimens from >200,000 patients and unaffected controls
Cancer Moonshot2020	Consortium	Phase 1: 20,000 cancer patients
Children's Hospital of Philadelphia Biorepository	Repository	Capacity for 8.6 million samples
China Kadoorie Biobank	Repository	>512,000 participants (general population, China). Genotyping data available for ~100,000.
CIMBA	Consortium	>15,000 BRCA1 carriers, >8,000 BRCA2 carriers
Clinical Sequencing Exploratory Research (CSER)	Consortium	~4,000 patients and healthy controls
DECIPHER	Repository	19,014 patients (international)
deCode Genetics	Organization	500,000 participants (international)
East London Genes & Health	Research Project	100,000 unaffected individuals (East London, Pakistani or Bangladeshi heritage)
Electronic Medical Records and Genomics (eMERGE) Network	Repository, Consortium, Research Project	55,028 patients
European Network for Genetic and Genomic Epidemiology (ENGAGE)	Research Project	80,000 GWAS scans, and DNA and serum/plasma from >600,000 individuals
Exome Aggregation Consortium (ExAC)	Consortium	60,706 individuals
GENIE/AACR	Data-Sharing Project	>17,000 cancer patients (international)
Genome Asia 100K	Consortium	Goal: 100,000 individuals (Asia)
Genomics England	Organization	Goal: 100,000 genomes from 70,000 individuals (rare disease & cancer patients, and their relatives)
GoT2D	Consortium, Data-Sharing Project	Multiple case-control cohorts
International Cancer Genome Consortium (ICGC)	Consortium	currently data from >16'000 samples
International Genomics of Alzheimer's Project (IGAP)	Consortium	40,000 patients with Alzheimer's disease
International Multiple Sclerosis Genetics (IMSG) Consortium	Consortium	Goal: >50,000 patients with MS
Kaiser Permanente: Genes, Environment, and Health (RPGEH)	Repository, Research Project	200,000 DNA samples (scaling up to 500,000)
Leiden Open Variation Database (LOVD)	Repository	>170,000 individuals
Million Veteran Program	Research Project	Goal: 1 million individuals; first 200,000 is complete.
MyCode® Community Health Initiative	Repository, Research Project	Goal: >250,000 patients
Precision Medicine Initiative	Research Project	Goal: >1 million participants, starting in 2016 (US)
Psychiatric Genomics Consortium (PGC)	Consortium	>170,000 subjects
Resilience Project	Research Project	589,306 individuals
Saudi Human Genome Program	Research Project	Goal: ~100,000 patients and controls (Saudi Arabia)
Scottish Genomes Partnership (SGP)	Research Project	>3,000 individuals (Scotland)
T2D-GENES	Consortium, Data-Sharing Project	10,000 patients and controls (five ethnicities); 600 individuals (Mexican American)
TBResist	Consortium	>2,600 samples
UK Biobank	Repository, Consortium, Research Project	500,000 individuals (age 40-69 years; UK)
UK10K	Research Project	10,000 participants (6,000 patients and 4,000 controls)
Vanderbilt's BioVU	Repository	>215,000 samples

National Medical Genome Projects and Cohorts (2018)

Medical Genomes

- Countries with active national medical genome projects
- Countries with some activity of medical genomics
- Countries planning medical genome projects

Cohorts

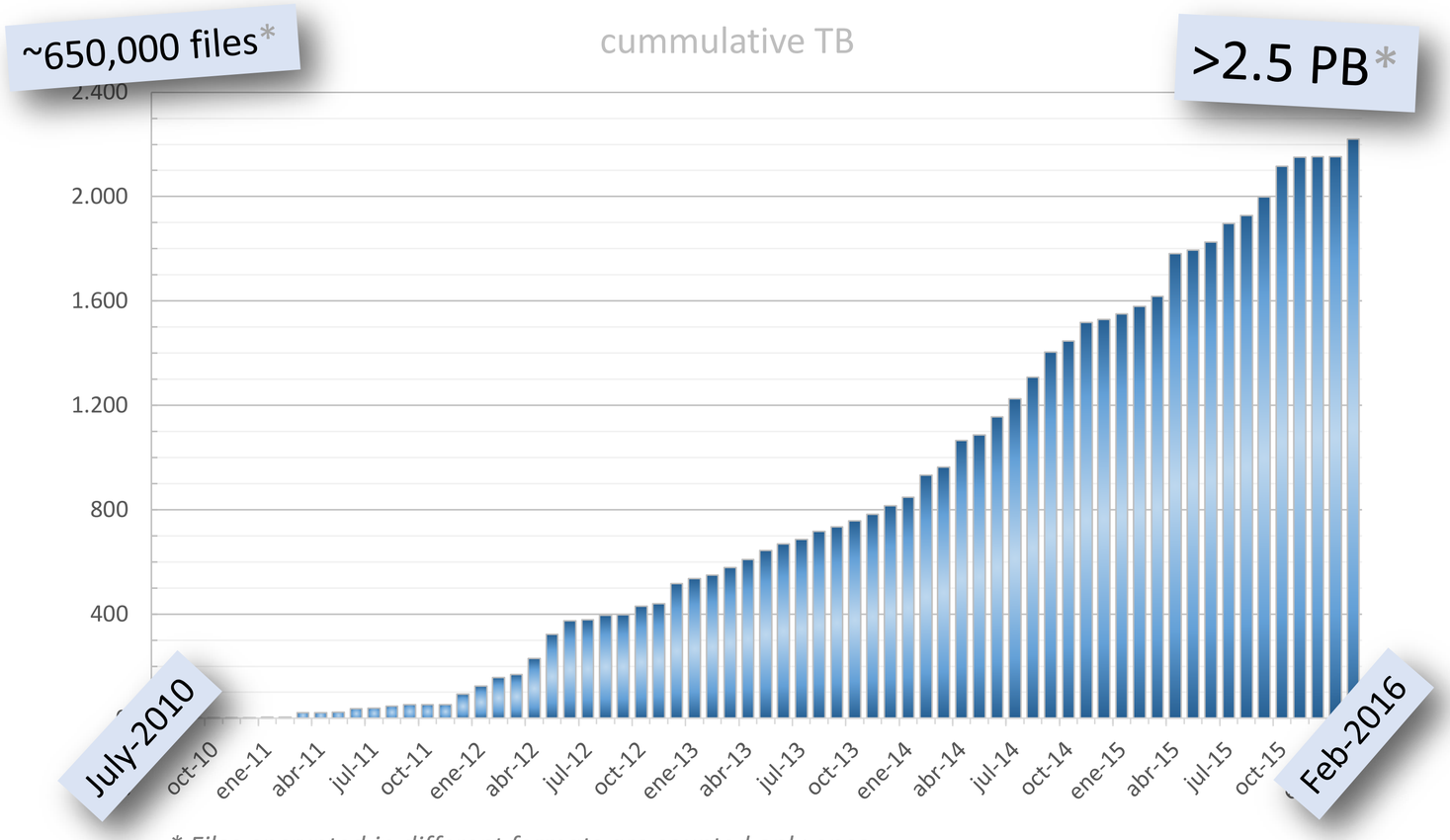
National cohorts > 100k genotyped or sequenced at least 25k

National cohorts > 100k people active collection now

Planning national cohorts > 100k

Growth of Genome Data Repositories: Example EGA

The EGA contains a growing amount of data



International Cancer Genome Consortium

IRDiRC INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

UK 10K

wellcome trust sanger institute

** Files encrypted in different formats are counted only once*

Source: Jordi Rambla (CRG Barcelona) 2016

Total worldwide sequencing capability?!

Deployed sequencers of major platforms

- deployed sequencers of major platforms (estimate end of 2021)
- "While 3 Exabases seems like a awful lot, it’s worth noting that this still isn’t enough to sequence every human born. Somewhere in the region of 4 children are born a second, our 100Gb/s wouldn’t even let us fully sequence one of them (at 30x)" - Nava Whiteford

Platform	Estimated Instruments	Runs/Week	Total Runs/Year	Run Yield (Tb)	Total Sequencing Capacity Tb/year	
ONT MinION	5501	3	858156	0.05	42907.8	
ONT GridION	782	3	121992	0.25	30498	
ONT PromethION 48	67	2	6968	14	97552	
Ilm Novaseq 6000	1485	3	231660	6	1389960	
Ilm NextSeq	5430	3	847080	0.36	304948.8	
Ilm Miseq/Mini/iSeq	12340	3	1925040	0.015	28875.6	
Ion Torrent	2220	14	1616160	0.05	80808	
PacBio	577	5	150020	0.03	4500.6	
MGI - Mid/Low	2000	3	312000	0.72	224640	
MGI -T7	10	5	2600	6	15600	
Total	30412		6071676		2220290.8	Tb/year
			2331523584	Samples	70.4	Gb/s
			73.9	Samples/s		

Estimated sequencing capability if everything would run continuously at full speed...

Source: 41J Blog - https://41j.com/blog/2022/09/global-sequencing-capability-notes/

Task: Reading up on Genome Technologies

- General NGS technologies
- count based vs. intensity based as principle
- bonus: dig deeper for some molecular-cytogenetic techniques:
  - SNP, aCGH arrays
  - SKY, M-FISH
  - chromosomal CGH

BIO392: Course Schedule

https://compbiozurich.org/courses/UZH-BIO392/

	Tue Sep 20	Wed Sep 21	Thu Sep 22	Fri Sep 23	Tue Sep 27	Wed Sep 28	Thu Sep 29	Fri Sep 30	Tue Oct 4	Wed Oct 5	Thu Oct 6	Fri Oct 7	Tue Oct 11	Wed Oct 12
09:00 - 10:00			Izaskun: Terminal / Unix / Files	Izaskun: File formats for human genetic variation / file handling		Michael lecture introduction to some resources, CNVs, Progenetix	Hangjia: Blast	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
10:00 - 11:00		Github exercise: create user specific directories & upload/edit test files using Markdown (Ziying)	Izaskun: Terminal / Unix / Files	Izaskun: File formats for human genetic variation / file handling		Task: Browse/explore genome resources and provide some notes (1-2 pages total) in a doc posted on Github (.md)	Hangjia: blast	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
11:00 - 12:00		Ziying: github desktop and terminal		Izaskun: File formats for human genetic variation / file handling			Hangija: Blast exercise	Max: STRs		Rahel: Sequence analysis	Rahel & Feifei: Survival analysis	Rahel & Feifei: Survival analysis		Exam
13:00 - 14:00	* Room information * Administrative - discuss times/days - exam	Ziying: Introduction to different interfaces eg atom, jupyter, pycharm (lecture), include R things	Izaskun: SIB online introduction to Unix	Izaskun: short project (1000 genomes), reading, literature	Recap W1; Q&A	Hangjia: Progenetix as tool for CNV frequencies etc.	Hangjia: Clinvar and Clingen	Max: STRs	Survival lecture, explanations of terms used etc, cancer classifications	Rahel: Survival analysis	Rahel & Feifei: Survival analysis	Discussion of Survival results (groups of groups)	Exam revision, Q&A	
14:00 - 15:00	Tina Siegenthaler: technical introduction (room, computer, accounts)		Izaskun: SIB online introduction to Unix	Izaskun: short project (1000 genomes), reading, literature	Literature (genome analysis techniques ...)			Max: STRs		Rahel	Rahel & Feifei: Survival analysis			
15:00 - 16:30	* explore course site * create Github accounts and forward to bio392@compbiozurich.org * Michael: short introductory lecture about genome variation			Izaskun: short project (1000 genomes), reading, literature	Genome technologies - brief notes about usage scenarios, pro & con			Max: STRs		Rahel	Rahel & Feifei: Survival analysis			