V SIMPOSIO GETHI | 18/19 noviembre de 2019

Ilustre Colegio Oficial de Médicos de Madrid. Aula Jiménez Díaz. Madrid

Tecnologías genómicas Next Generation Sequencing: amplicones versus captura híbrida

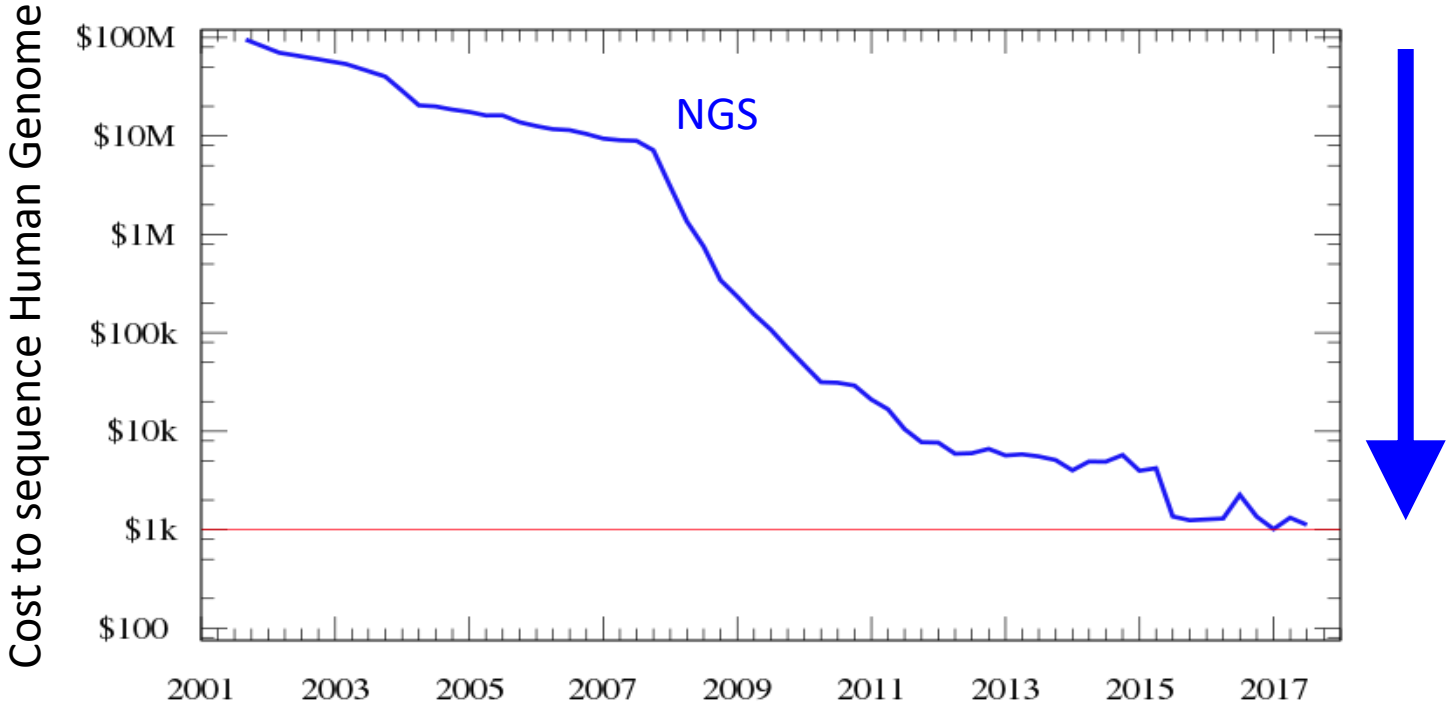
Cristina Rodríguez-Antona

Centro Nacional de Investigaciones Oncológicas
Programa de Genética del Cáncer Humano
crodriguez@cnio.es

Next-Generation Sequencing (NGS)

Fundamentally different approach to sequencing that has made a revolution in genomic medicine

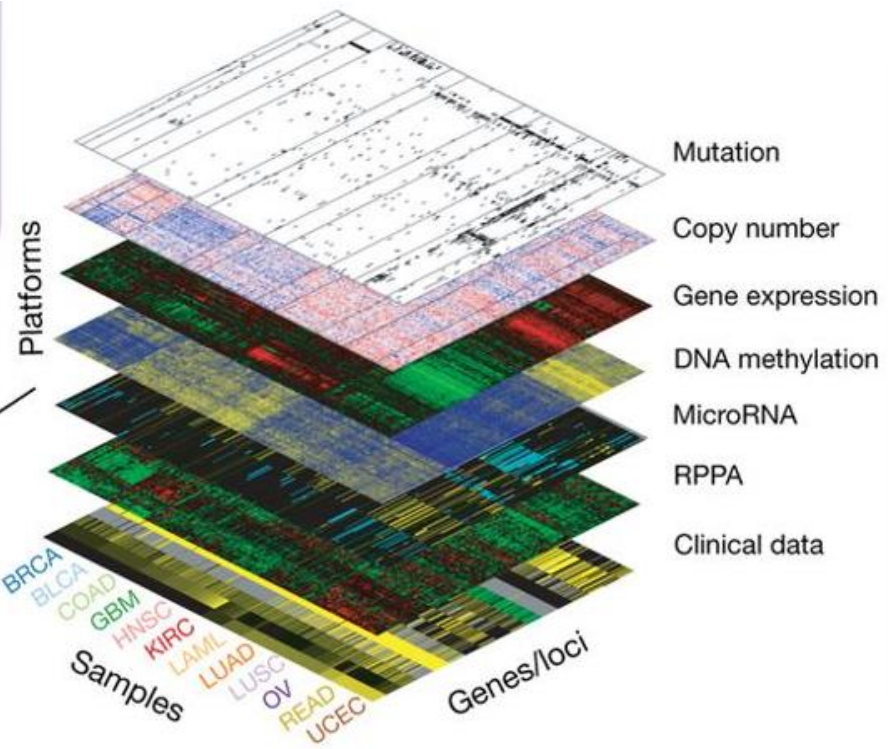
- 1 human genome
- 3000 million $
- 13 years work

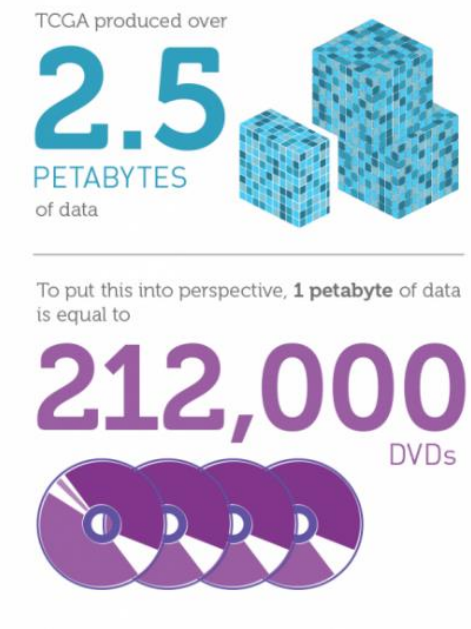


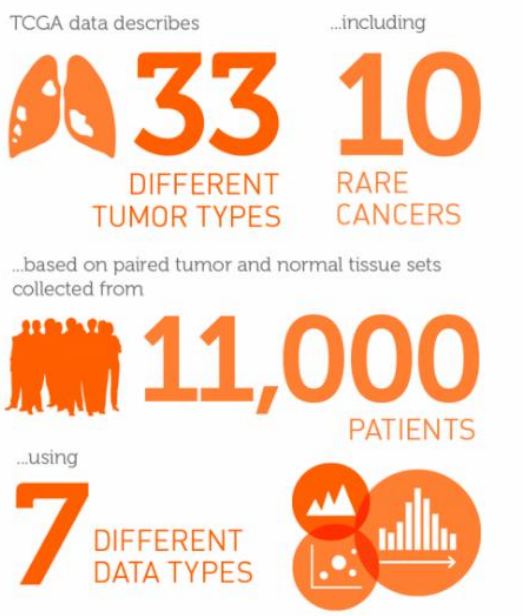
- 1 human genome
- 1000$
- 1 day work

Reference International **Research Projects** in cancer

The Cancer Genome Atlas 2005 - 2018





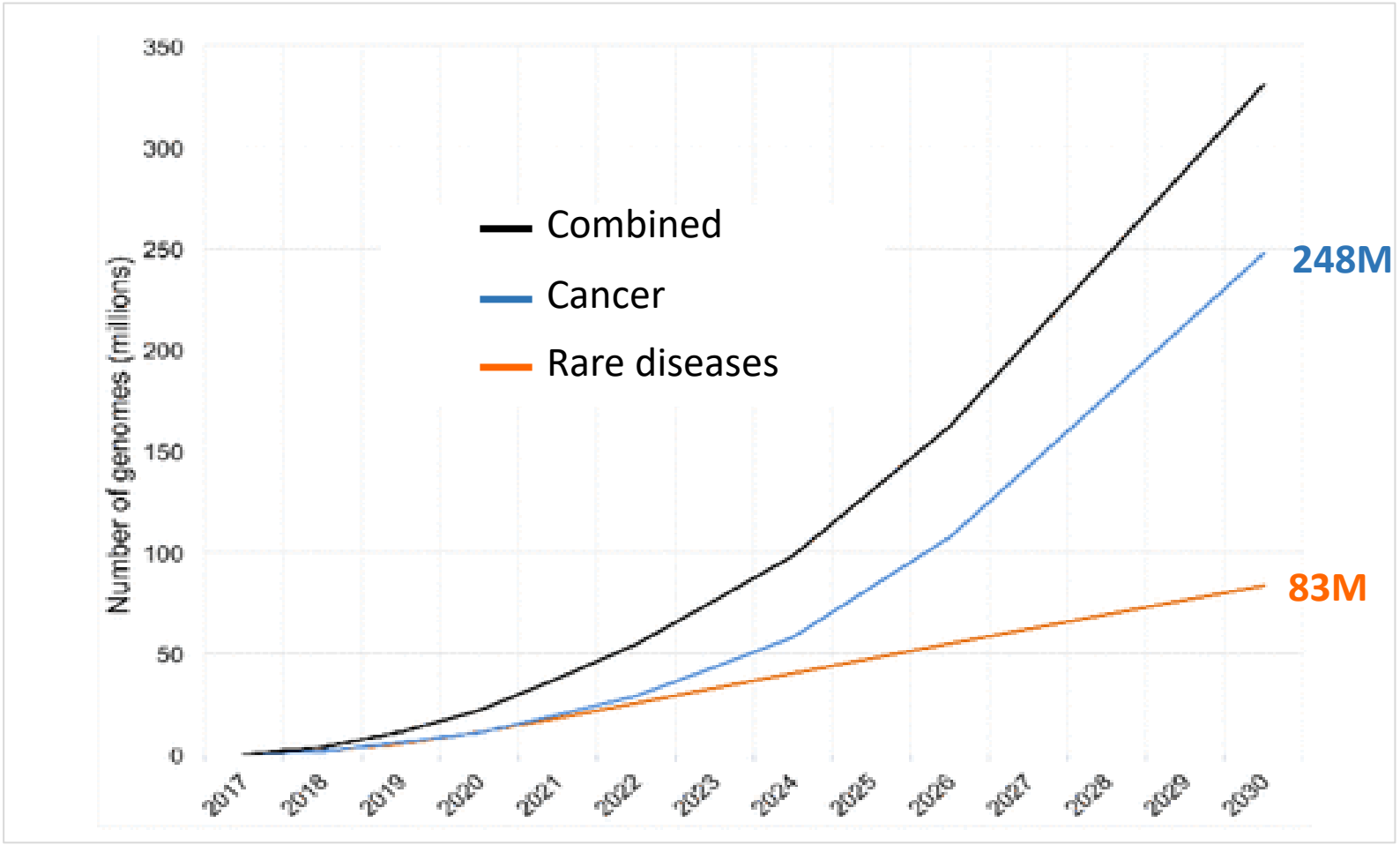


- Adrenocortical Carcinoma
- Adult ALL (B-cell and T-Cell)
- Anaplastic Thyroid
- Cholangiocarcinoma
- Chromophobe kidney
- High Risk MDS (del 5q- cases)
- Mesothelioma
- Pheochromocytoma
- Testicular Germ Cell
- Thymoma
- Uterine Carcinosarcoma
- Sarcomas

Daily/ weekly consulted in cancer research lab

Application of WES & WGS in **Healthcare**

Exponential growth in genome sequences from rare diseases and cancer genomics



Genomics in healthcare: GA4GH looks to 2022. doi: https://doi.org/10.1101/203554

National Genomic Initiatives

The 100,000 Genomes Project

Sequence 100,000 genomes from NHS patients with rare disease or cancer.

The 100,000th sequence in December 2018.

Actionable findings for **1 in 4-5** rare disease patients, **50%** cancer cases contain the potential for a therapy or a clinical trial.

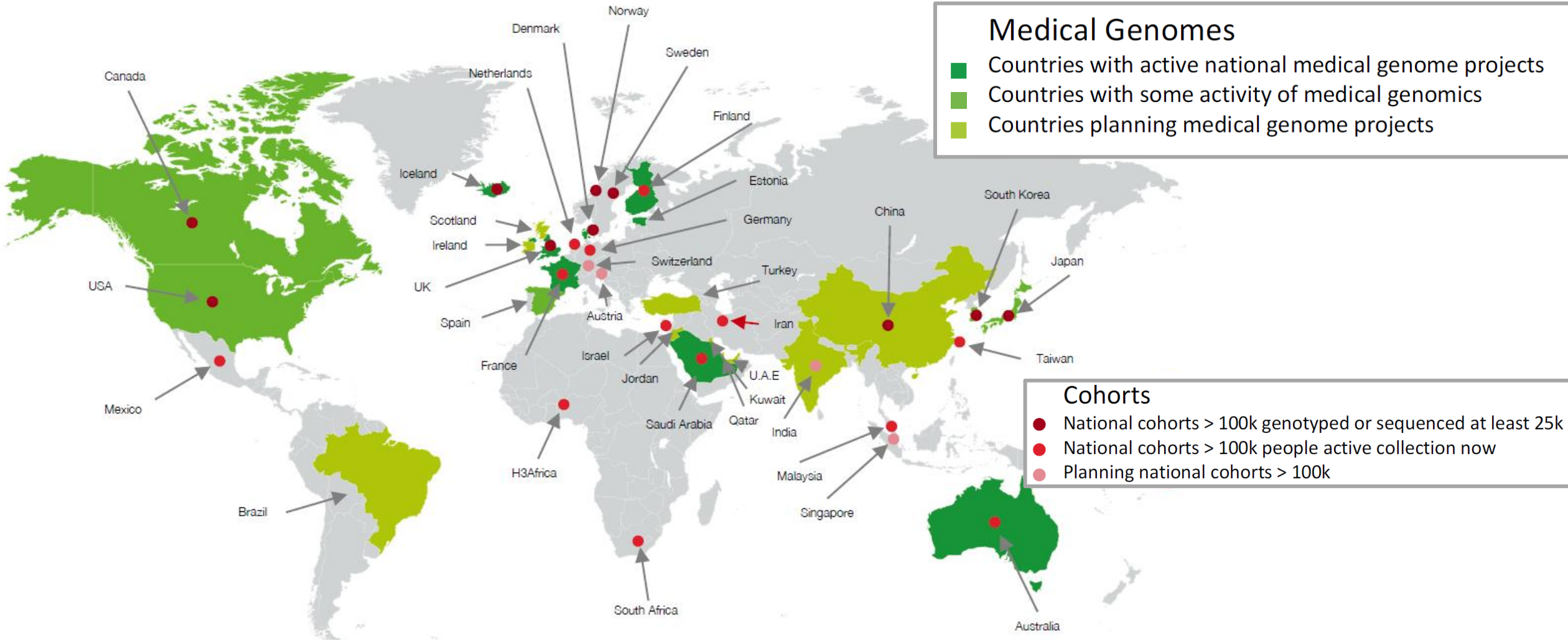
Precision Medicine Initiative

>1 million volunteers, provide genetic data, biological samples, information about their health.

Short-term goal: expand precision medicine in cancer research.

Long-term goal: bring precision medicine to all areas of health and healthcare

National **Medical Genome** Projects and Cohorts

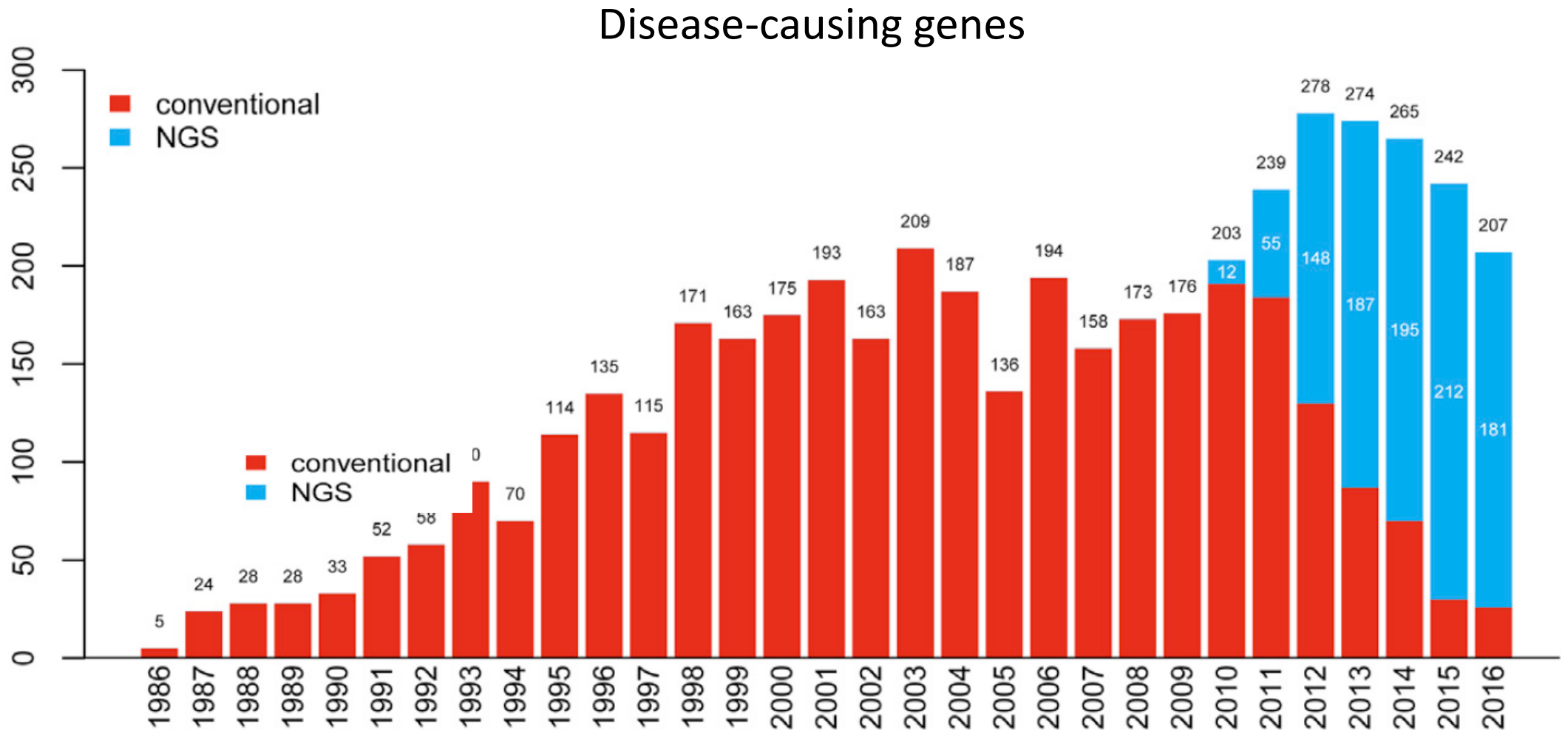


Data jointly gathered by EMBL-EBI and GA4GH (Sept 2018)

Global Alliance for Genomics & Health
Collaborate. Innovate. Accelerate.

EMBL-EBI

NGS for identification **genetic markers for healthcare**



AJHG 2017;100:695

Human Genome

- 3.200 millions of base pairs
- 20.000 genes (1-2% of genome)
- >200.000 coding exons

- >**4 M** single nucleotide variants per individual
- >**20.000** structural variants per individual
- **10-20 "loss-of-function"** variants in healthy individual (no phenotype)

Discovery causal gene

nature genetics

Letter | Published: 19 June 2011

Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma

One single shared gene segregated with disease: *MAX*

Germline Whole Exome Sequencing of 3 unrelated patients with Familial PCC

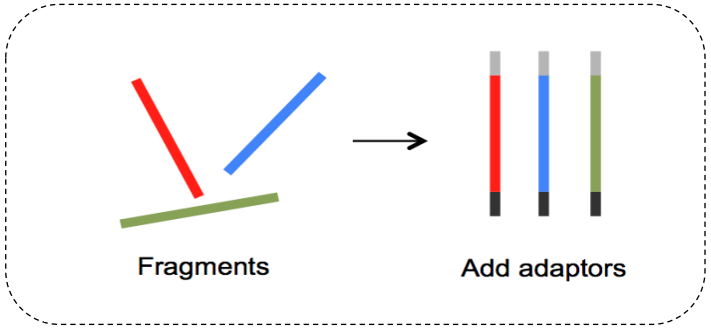
Sample number	Total number of SNSs	Heterozygous SNSs	SNSs after HapMap control filtering	SNSs after dbSNP130 filtering	SNSs after removing intronic and intergenic variants	SNSs affecting the same gene in the three samples	SNSs in coding regions (not in UTR)	SNSs after removing synonymous, deep<7, and Phred_qual<20	SNSs after removing entries in additional databases[a]
924	95,100	28,292	9,088	2,911	763	41[b]	17	9	5
3037	92,855	26,859	7,884	3,066	789				
3121	89,784	26,232	8,292	3,676	743				

Comino-Mendez, Nat Genet. 2011;43:663

Best NGS technology?

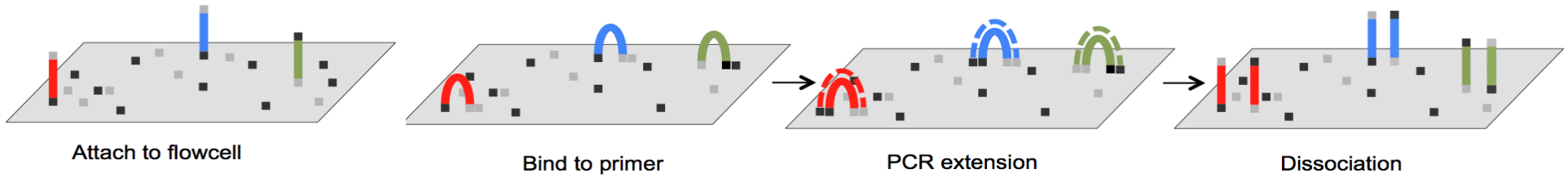
Next Generation Sequencing steps

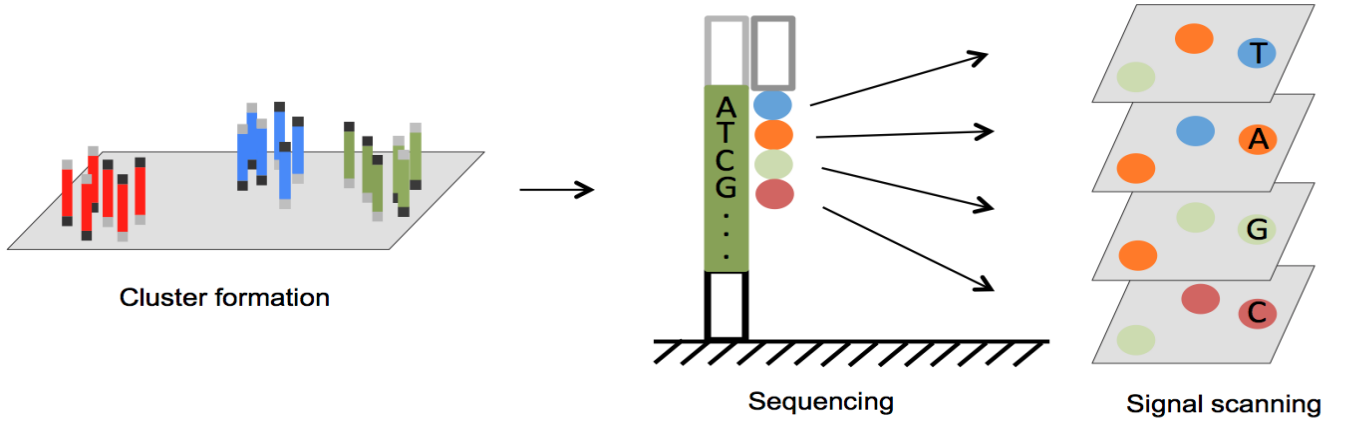
1. Samples-Library preparation



2. Massive sequencing

Illumina



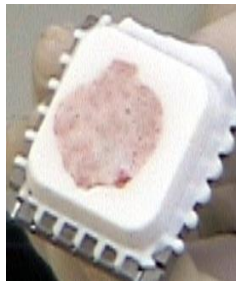


3. Bioinformatic analysis

Raw data → **Annotated variants**

Samples for NGS

blood



frozen tumor

FFPE tumor

plasma

stool

urine

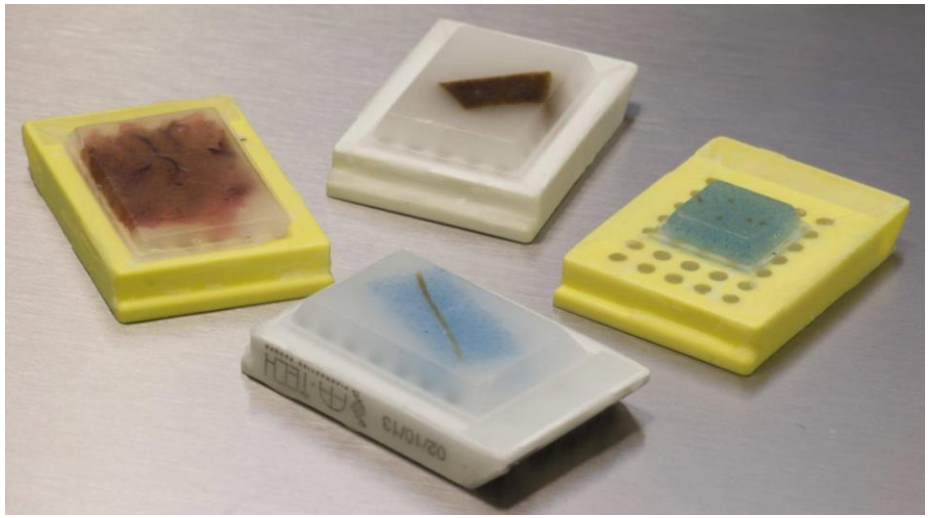
(…)

DNA or RNA isolation

Different strategies

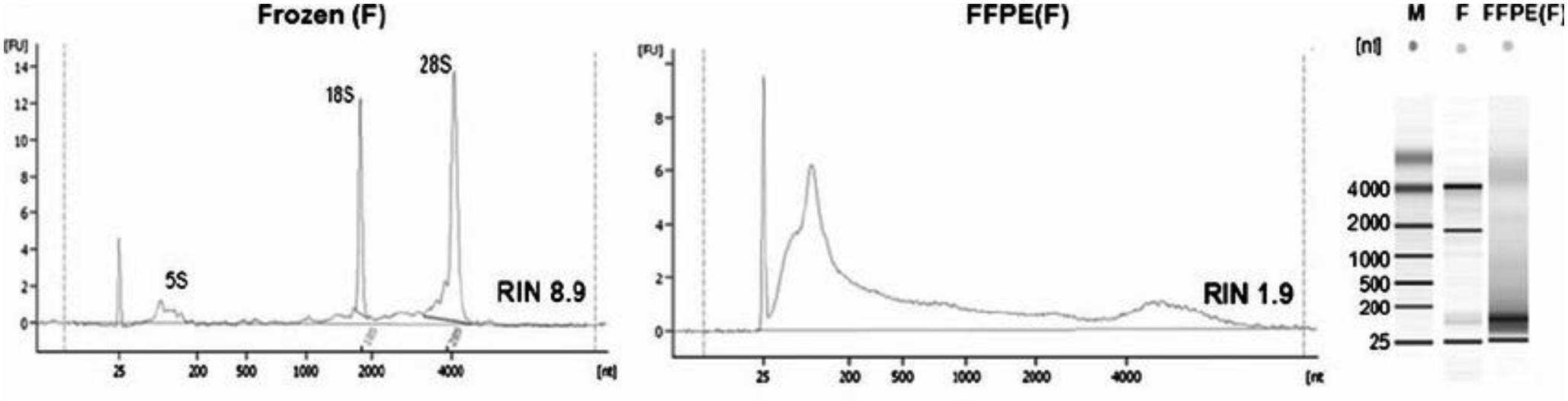
Library & NGS

Samples for NGS



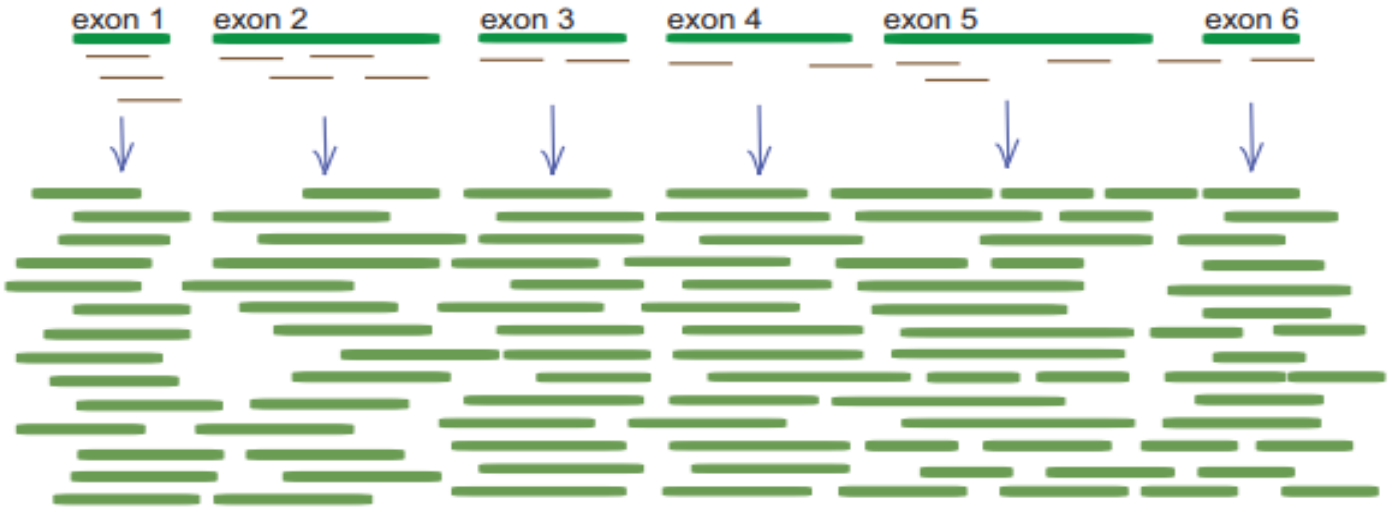
All techniques in all samples? → NO

- Amount tissue
- Costs
- Quality
- Limitations (SV)



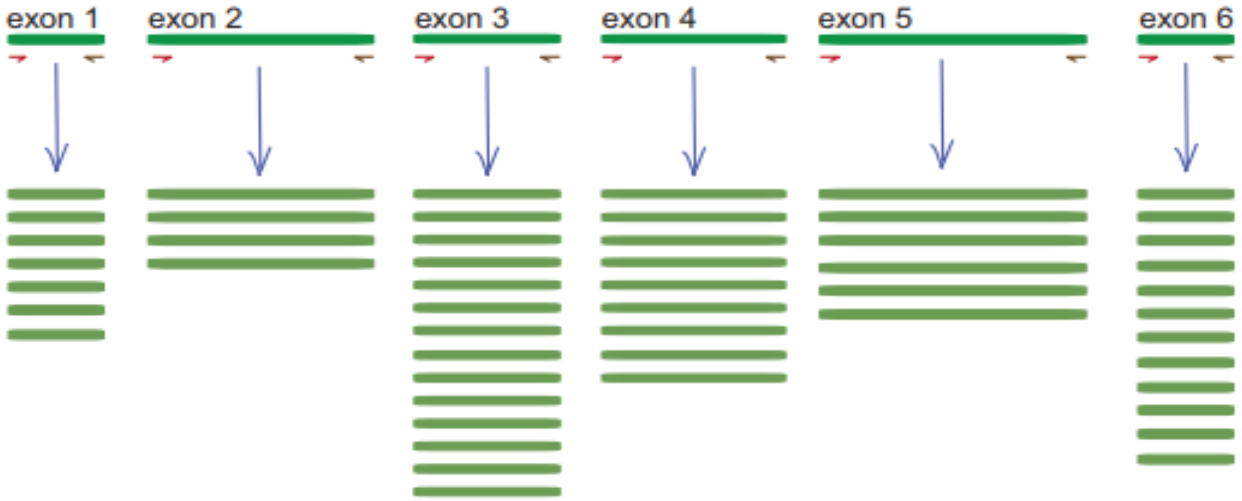
Library preparation

Hybridisation



All fragments identical. Duplicates cannot be distinguished from unique products. Assay artefacts cannot be distinguished.

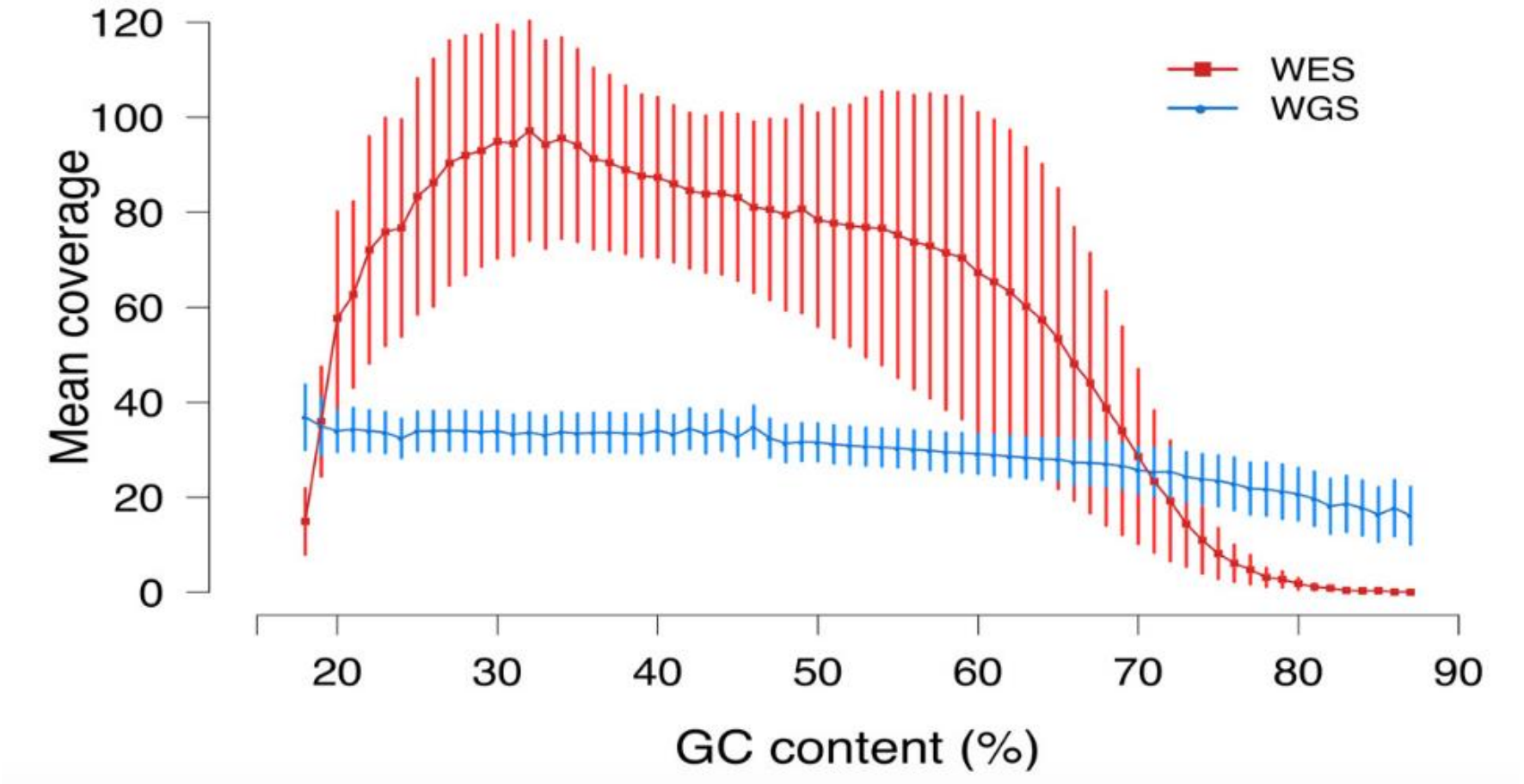
Amplicon



Primer competition and preferential amplification will lead to non-uniform enrichment.

Library preparation

Coverage over GC content in WGS and WES

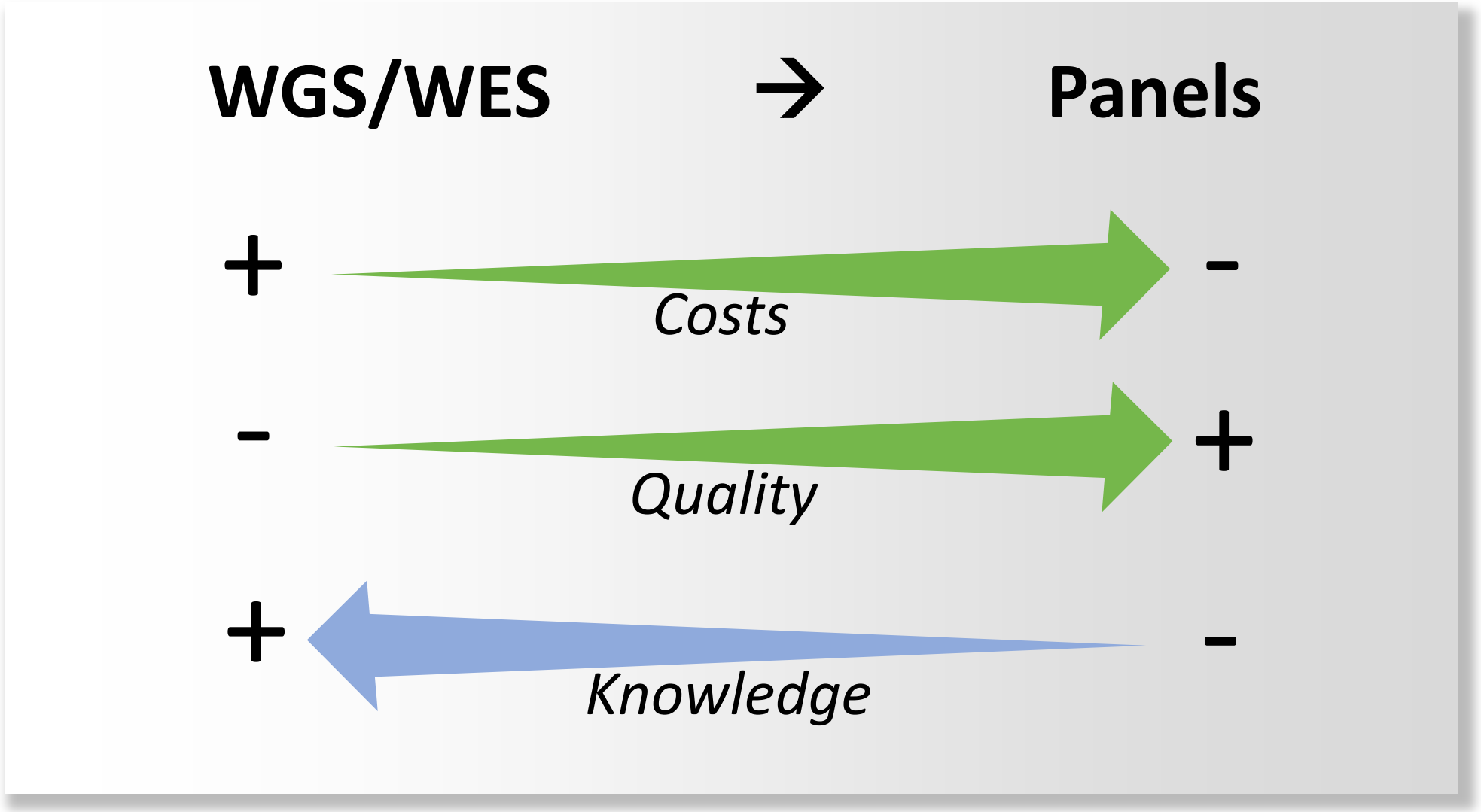


Björn et al. DOI: 10.4172/2469-9853.1000154

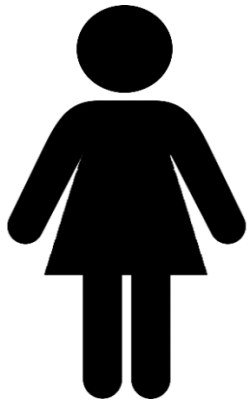
Sequencing platforms

Feature	HiSeq2500 - Highoutput	HiSeq2500 – Rapid mode	MiSeq	PacBio RSII
Number of reads	150-180M/lane	100-150M/lane	12-15M (v2) 20-25M (v3)	50-80K/SMRT cell
Read length	2 x 100 bp	2 x 150 bp	2 x 300 bp (v3)	~ 10-20 kb
Yield per lane (PF data)	up to 35 Gb	up to 45Gb	up to 15 Gb	up to 0.4 Gb
Instrument Time	~12-14 days	~2 days	~2 days	~2 hours
Pricing per Gb	$59 (PE100)	$53 (PE150)	$108 (PE300)	$697

Choosing the best technology



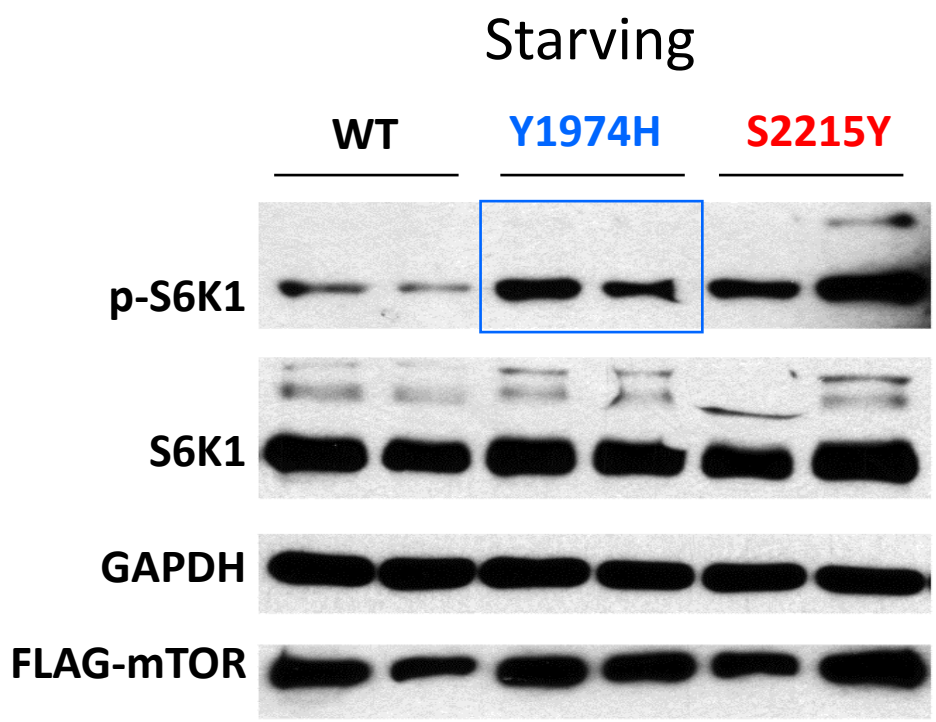
Variants of Unknown Significance (VUS)



57 y woman

- **Metastatic ccRCC** (bone L3-5, liver)
- Nephrectomy (2007)
- **Sunitinib** > 2 cycles PD; **Sorafenib** > 2 cycles PD
- **Temsirolimus** 2007 **> prolonged response**
- Continuos treatment (good tolerance)
- **2015** skull mtx > resection + radiotherapy
- Currently, disease free with temsirolimus

Functional assay



Rodriguez-Moreno, J Natl Compr Canc Netw. 2017;15:1310

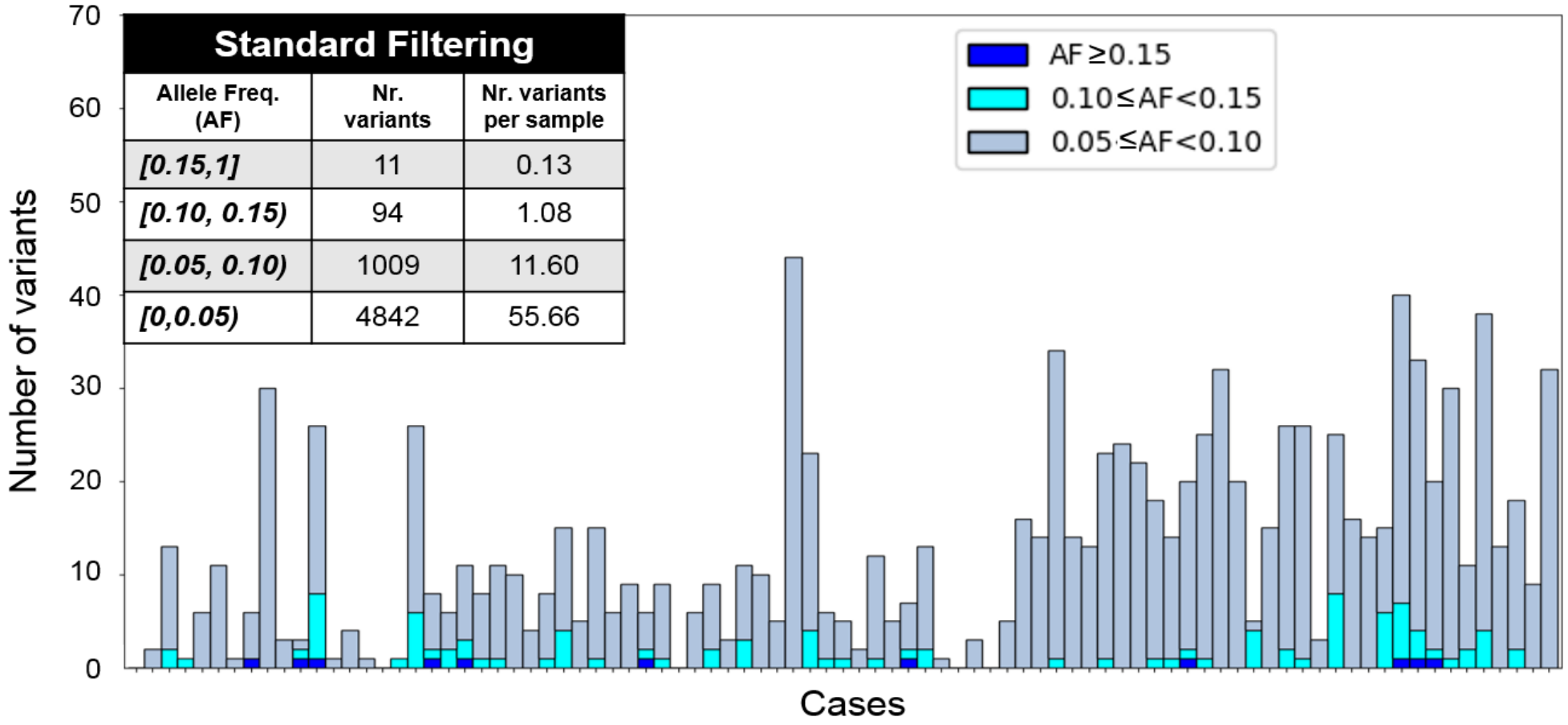
HOSPITALES

J.F. Rodríguez
J. García-Donas
I. Durán

Bioinformatic analysis

***MTOR, TSC1* or *TSC2* mutations in RCC**

Standard Filtering		
Allele Freq. (AF)	Nr. variants	Nr. variants per sample
[0.15,1]	11	0.13
[0.10, 0.15)	94	1.08
[0.05, 0.10)	1009	11.60
[0,0.05)	4842	55.66



J. García-Donas

HOSPITALES

Roldan-Romero, Int J Cancer. 2019 in press

Conclusions

- Large genomic national projects exist (cancer and rare diseases), with exponentialy increase of WGS/WES in medicine.

- Technical approaches suitable for different sample types exist (e.g. FFPE, plasma) and additional resources are evolving (e.g. long reads). However, challenges still exist (e.g. in homogeneity of coverage, bioinformatic analysis, variant interpretation, incidental findings…)

- Next Generation Sequencing is an enormous force transforming conventional medicine into a more efficient and safer personalized medicine

Acknowledgements

cnïo
stop cancer

Patients
&
Hospitals

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD

Obra Social
Fundación "la Caixa"

Pfizer

SOGUG

HM
HOSPITALES