Medicina personalizada en el sistema público de salud: la experiencia en Andalucía

Joaquín Dopazo

Area de Bioinformática Clínica, Fundación Progreso y Salud, Functional Genomics Node, (INB-ELIXIR-es), Bioinformatics in Rare Diseases (BiER-CIBERER), Sevilla, Spain.

http://www.clinbioinfosspa.es http://www.babelomics.org

XI Conferencia Anual de las Plataformas Tecnológicas de Investigación Biomédica, Barcelona, 6 marzo, 2018

The Personalized Medicine Plan of the Andalusian community

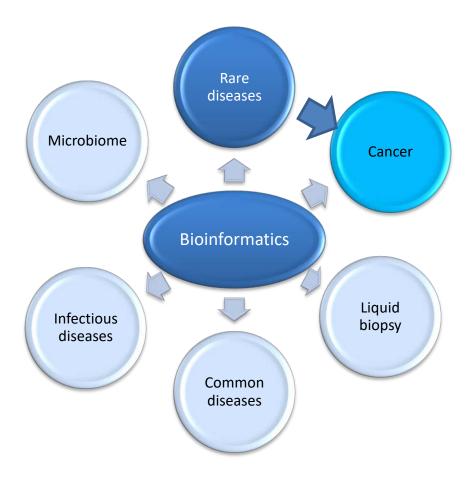
The Program of Personalized Medicine of the Andalusian Community aims to use patient genomic data for precision diagnostic and treatment recommendation, with the long-term idea of converting the whole health system into a prospective clinical study. Andalusia is the third largest region in Europe with a population of 8.5 M inhabitants



Andalusia is the <u>largest health</u> <u>system in Europe</u> having 8.5M people under a <u>unique</u>, <u>universal</u> <u>electronic health record</u>.

Bioinformatics as a fundamental piece of the Personalized Medicine Plan of Andalucía

The Bioinformatics Area, created in June 2016 in the Fundación Progreso y Salud, has as main goal supporting the Program of Personalized Medicine of the Andalusian Community by facilitating the use of genomic data for precision diagnostic and treatment recommendation. A pilot project in rare diseases will be followed of the progressive implementation in cancer, common diseases, infectious diseases and other emergent, clinically relevant aspects, such as liquid biopsy, microbiome, nutrigenomics.



Previous initiatives The MGP and CIBERER Initiatives to sequence rare disease patient exomes



Sample providers

Sequencing platforms

Data analysis

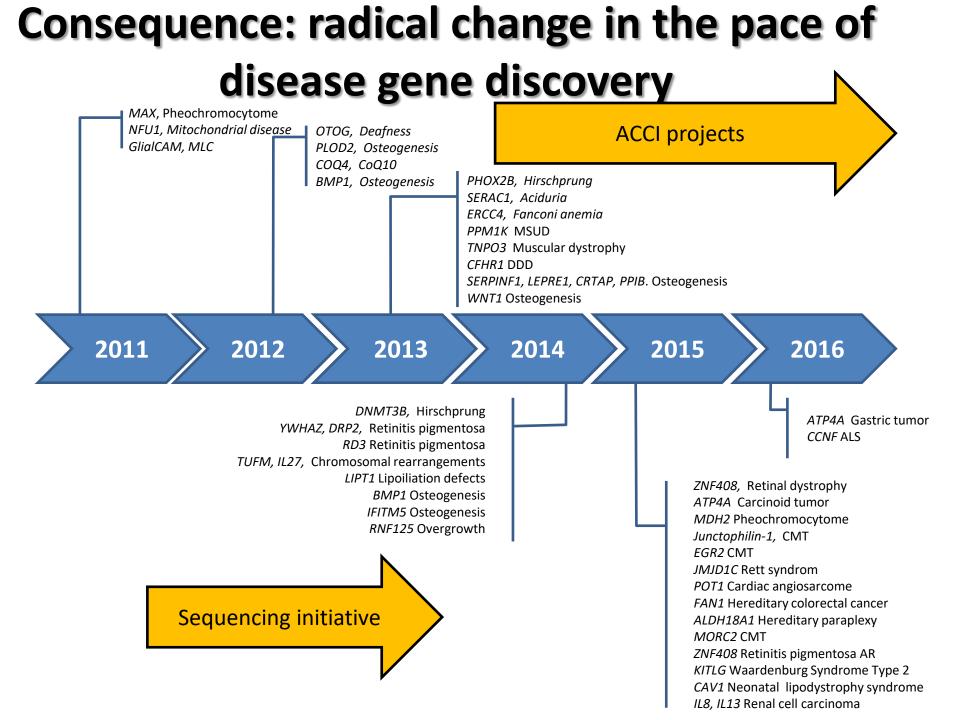
Diseases with

- Unknown causal genes
- No mutations in known genes

Search for:

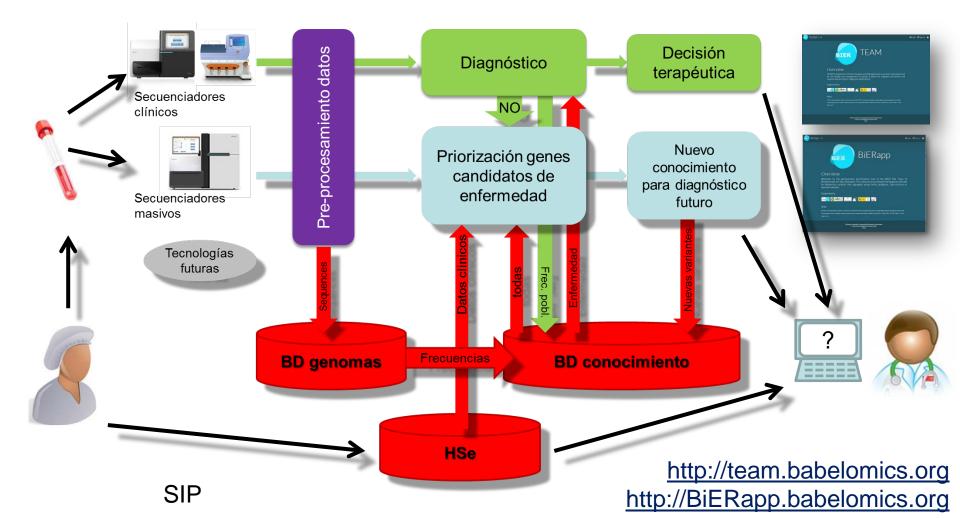
- New disease genes
- Known genes with unknown modifier genes
- Susceptibility genes

A total of 1044 exomes of 300 healthy controls and patients of more than 30 diseases were sequenced between 2012 and 2013.



CIBERER initiative for diagnosis and biomarker discovery using massive sequencing

Ongoing (second phase) Project with the collaboration of seven hospitals: La Paz, FJD, Ramón y Cajal, CBM (Madrid), Virgen del Rocío (Sevilla), Hospital del Mar (Barcelona), HU La Fe (Valencia)., **within the context of CIBERER**



The CSVS: a crowdsourcing project

		B	IE	R		2S • s	V S	5													
	Ov	ervie	w																		
	inform commo of pri unrel	mation ab nunity. It i ioritizing	out the s useful candidat nish indi	variability for filterin e disease viduals. W	oanish Varian of the Spani g polymorphi genes, CSVS /e accept sub	sh popu isms an curren	ulation d local tly stor	to the variations res info	scient ons in rmati	ific/me the pro on on 1	dical cess .582										
	Supp	orted by																			
	cheve	a cibe	rer 🧐	Tenate Pages	er 🛋 💰	755.	E eli)	Úr.													
	200																				
		: ets application r vents, so only m			HTML5 standard ar y supported, these in												l	Ì			
	CSVS va	: ets application r vents, so only m		rowers are full	y supported, these in	elude Ove														1000G M	W [phase
	CSVS va	: ets application r vents, so only m		rowers are full		elude Ove											EUR	ALL		1000G MJ South AS	
C. M omotoma Locabore	CSVS va	: ets application r vents, so only m		rowers are full	y supported, these in	elude Ove											EUW C2NC	ALL 0.263			
Chi oneonal Localon 111-146009,211-130000	CSVS-w Compor Opera 2	: et application r entra, vo only m id4.	CIA	csvs: o CSV5: o RET	y supported, these in reated by Clinical Bio Fundacide Progress 2013-2017 rt329902197	elude Cive informatics y Salud \$75	ine 36+, Fi Area	retax 32+. 1	ie 10+, 54	elari 7+ ani 0.1997	C003	0.003	0.350	6430	0.240	0440	C.200 C.400	0.263	AME 0.248 0.438	South AS 0.300 0.410	0.482 0.305
2:1-100000,2:1-100000	CSVS w Compar Opera 2 50	: eta application r nenku, so only m 14+. 43572513 43572533	CiA CiA	CSV5:0 EET EET	y supported, these in reated by Clinical Bio Fundación Progress 2013-2017	elude Ove informatics y Solution 575 508		refox 32+. 1 18	e 10+, 54 0 0	0.997 0.124	C003 C078	0.076	0.360	6.430 0.400	0.210	0.440	6.2%	0.2+3	AME 0.248	South AS	0.482 0.305
	CSVS w Compar Opera 2 30 30 30	: et application r nerks, vo only m lde. 43572513 43572513 43572513	CIA CIA CIA CIA	CSV3:0 EST BET BET	y supported, these in reated by Clinical Bio Fundacide Progress 2013-2017 rt329902197	eluse Ove eformatics y Soluti 575 508 575	ine 36+, Fi Area	refox 32+, 1 18 0	ie 10+, 54	elari 7+ ani 0.1997	C003 C07a C07a	0.076					C.200 C.400	0.263	AME 0.248 0.438	South AS 0.300 0.410	0.482 0.305
1:1-100000,2:1-100000 Gene:	CSVS w Compar Opera 2 50	: eta application r nenku, so only m 14+. 43572513 43572533	CiA CiA	CSV5:0 EET EET	y supported, these in reated by Clinical Bio Fundacide Progress 2013-2017 rt329902197	elude Ove informatics y Solution 575 508	ene 36+, Fi Area 2 52	refox 32+. 1 18	e 10+, 54 0 0	0.997 0.124	C003 C078	0.076					C.200 C.400	0.263	AME 0.248 0.438	South AS 0.300 0.410	0.482 0.305
1:1-100000,7:1-100000 Gene:	CSVS w Compar Opera 2 30 30 30	: et application r nerks, vo only m lde. 43572513 43572513 43572513	CIA CIA CIA CIA	CSV3:0 EST BET BET	y supported, these in reated by Clinical Bio Fundacide Progress 2013-2017 rt329902197	eluse Ove eformatics y Soluti 575 508 575	rne 36+, Fr Area 2 52 3	refox 32+, 1 18 0	0 0 0	0.997 0.124 0.197	C003 C07a C07a	0.076					0.280 0.400 0.400	0.263	AME 0.248 0.438 0.438	South AS 0.300 0.410 0.420	0.482 0.305
1:1-100000,2:1-100000 Gene:	CSVS w Compar Opera 2 30 30 30 30 30	: et application r renta, so only m (4- 43572511 43572532 43545828 43545828	CIA CIA CIA CITO CIT	CSV3: o EET EET EET	y supported, these in restead by Clinical Bio Fundacide Progress 2013-2017 rt30900297 rt32957460	eluse Ove informatics y Solution 575 575 575 575	ame 36+, Fa Area 2 52 3 3	1 38 0 0	0 0 0 0	0.997 0.124 0.197 0.197	C003 C078 C003 C003	0.076 0.003 0.003				0.330	C280 C400 C400	0.263	AME 0.248 0.438 0.438	South AS 0.300 0.430 0.430 -	0.482 0.305
111-198000, 2:1-190000 Gene: 827	CSVS w Compar Opera 2 30 30 30 30 30	: eta poplication n reneta, so only m 445572511 43572525 43595808 43595808 43595808	CNA GNA CNA CNA CNA CNT	CSVILO CSVILO RET RET RET RET	valgeorted, these in rested by Clinical Bio Fundación Progress 2013-2017 rt12267460	eluse Ove informatics y Solid 575 501 575 575 474	Area 2 1 3 73	1 18 0 0 33	0 0 0 0 0	0.597 0.124 0.197 0.193	C003 C038 C003 C003 C003 C119	0.076 0.003 0.003 0.139				0.330	C280 C400 C400	0.263	AME 0.248 0.438 0.438	South AS 0.300 0.430 0.430 -	0.482 0.305
In 1 - 1 moor, 2: 1 - 3 moore Gene: #E7 Subpopulations ♀ □ ✔ MCP (267 healthy controld)	CSVS w Compar Opera 2 30 30 30 30 30 30 30 30 30 30 30 30	: eta poplication riereta, so orly m reteta, so orly m 445-72513 43572533 43575838 43575838 43575838 43575838	CIA GIA GIA CTD- CIT IT-	CSVE o EET EET EET EET EET EET	valeted by Clinical Dir Portical Dir Portical Dir 2013-2017 11320000097 113228-7460	eluae Ove eformatics \$75 \$58 \$75 \$75 \$75 \$75 \$75 \$75 \$75 \$75	Area 2 52 1 1 71 71	1 38 0 33 33	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.997 0.124 0.197 0.891 0.891	C.003 C.078 C.003 C.003 C.003 C.003 C.003	0.076 0.003 0.003 0.139 0.139	0.350	0,400	0.270	0.330	C.2000 C.4000 C.4000	0.2+3 0.420 0.3+5	AME 0.248 0.438 0.418	South AS 0.300 0.430 0.430	C 487 C 482 C 385 C 385 C 385
In 1-1 mood, 21 1-1 mood Gene: #ET Subpopulations CO CO MCP (287 healthy	CSVS w Compar Opera 2 50 50 50 50 50 50 50	43572512 43572512 43572512 4357583 4359583 43595838 43595837 43595837 4359584	CVA GVA GVA CTG- CVT T- AG	CVVI o CVVI o RET RET RET RET RET RET	valeted by Clinical Dir Portical Dir Portical Dir 2013-2017 11320000097 113228-7460	eluae Crve enformatics y Salad 575 575 575 474 474 474	Area 2 52 1 71 71 74	1 18 0 33 33 156	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.997 0.124 0.197 0.197 0.891 0.881 0.666	C003 C07e C003 C003 C003 C003 C0139 C0139 C0139	0.076 0.003 0.003 0.137 0.139 0.139	0.350	0,400	0.270	0.330	C.2000 C.4000 C.4000	0.2+3 0.420 0.3+5	AME 0.248 0.438 0.418	South AS 0.300 0.430 0.430	C 482 0.305 0.305

http://csvs.babelomics.org/

Allelic population frequencies obtained from 1,600 exomes are currently available in CSVS Scenario: Sequencing projects of healthy population are <u>expensive</u> and funding bodies are <u>reluctant</u> to fund them CSVS Aim: To offer increasingly accurate information on variant frequencies characteristic of Spanish population. CSVS Main use: Frequency-based filtering of candidate variants Main data source: Sequencing projects of individual researchers (CIBERER and others)

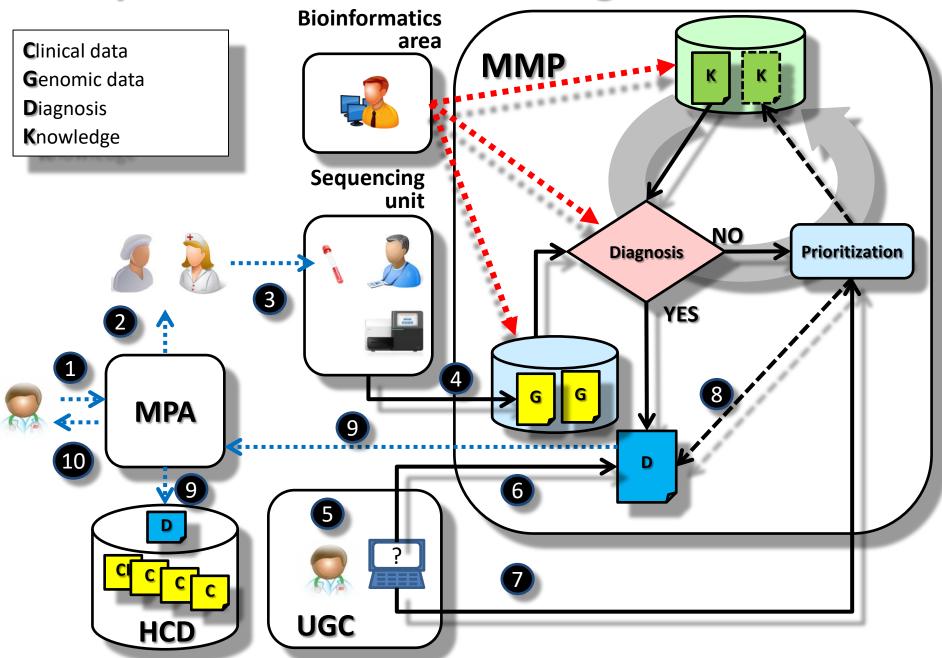
Problem: Most of the contributions are patient exomes

Idea: Patients of disease A can be considered healthy **pseudo-controls** for disease B (providing no common genetic background exist between A and B) Beacon: CSVS has a Beacon server

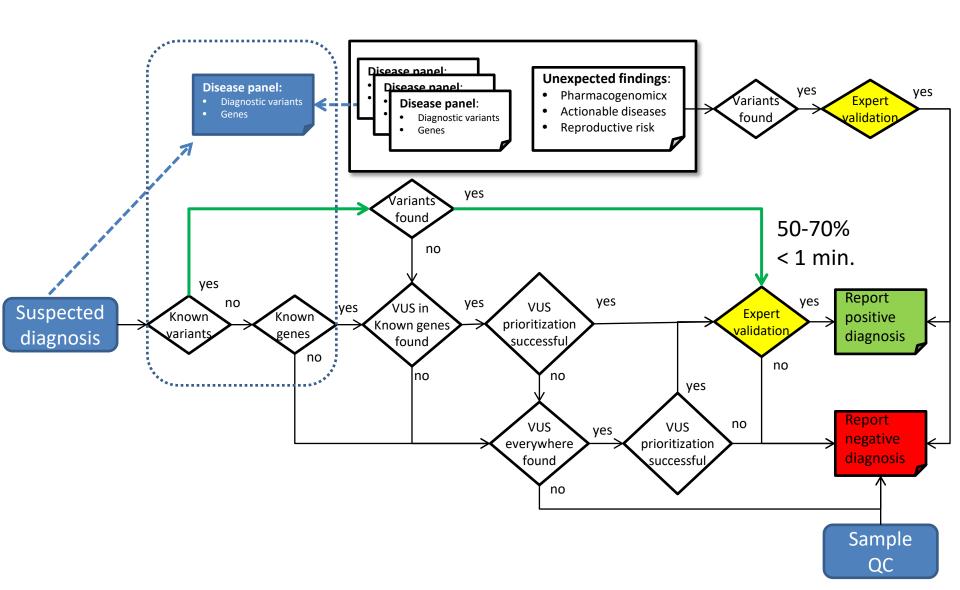
2017 - First phase of the Personalized Medicine Plan of Andalucía: rare diseases Objectives:

- Implementation of a pilot project to use <u>massive sequencing</u> for <u>high</u> precision diagnostic of rare diseases within the SAS
- Change in the current protocols to <u>introduce of massive sequencing</u> as part of the <u>clinical practice</u>: exome sequencing is cheaper than more than three conventional Sanger sequencing tests and produces much more clinically useful information
- Implementation of a database of genomic data linked to patient's eHR
- <u>Scalability</u> to the dimensions of a whole health system (Andalusia is the third largest region in Europe with >8.5 million inhabitants)
- Setting the foundations for the <u>prospective healthcare</u> paradigm by using the genomic database for prospective clinical studies (e.g. search for adverse response biomarkers...)

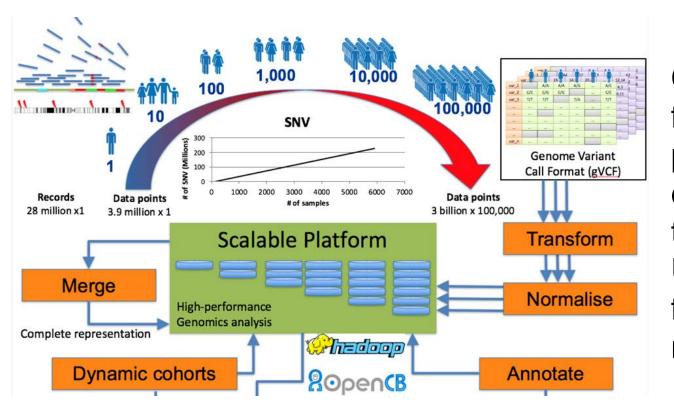
First phase: rare disease diagnosis in the SAS



General diagnosis protocol



Back end: OpenCGA, a scalable storage and genomic data management platform





In collaboration with Genomics England (GEL)

Currently, the fastest and more powerful genomic database engine in the world. Used in the GEL for genomic data management

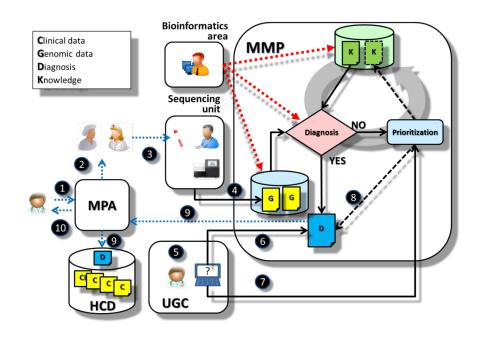
Extensive capabilities to query across genotype and phenotype relationships

https://github.com/opencb/opencga

GDPR compliance

The system has been designed in a way that is compliant with EU and Spanish General Data Protection Regulation

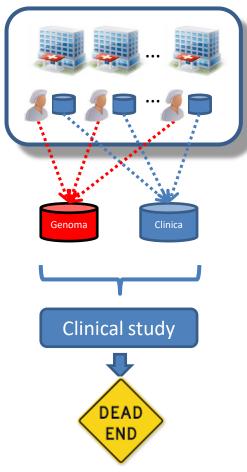
- Clinicians requesting for a genomic diagnostic have access to eHR and get the result of the test.
- Geneticists have access to eHR and can query the genomic data (but never extract them)
- IT have access to deidentified genomic data and no to eHR.



Genomic + clinical data = treasury

- Database of patients with prospective clinical information. Patients sequenced:
 - will have different responses to treatments
 - can have other diseases in the future
 - Today's undiagnosed patients can have a future diagnostic
- More <u>opportunities</u> for <u>precision genomic diagnostic</u> <u>without</u> the necessity of <u>genomic data transfer</u>:
 - Recurrent occurrence of variants in patients of the same disease favors finding new diagnostic disease variants and biomarkers.
 - Variants in patients of unrelated diseases can be taken as population background genomic variability to discard potential disease causal variants.
- Preventive medicine:
 - Prospective discovery of pharmacogenomic relevant variants
 - Prospective discovery of new disease risk variants

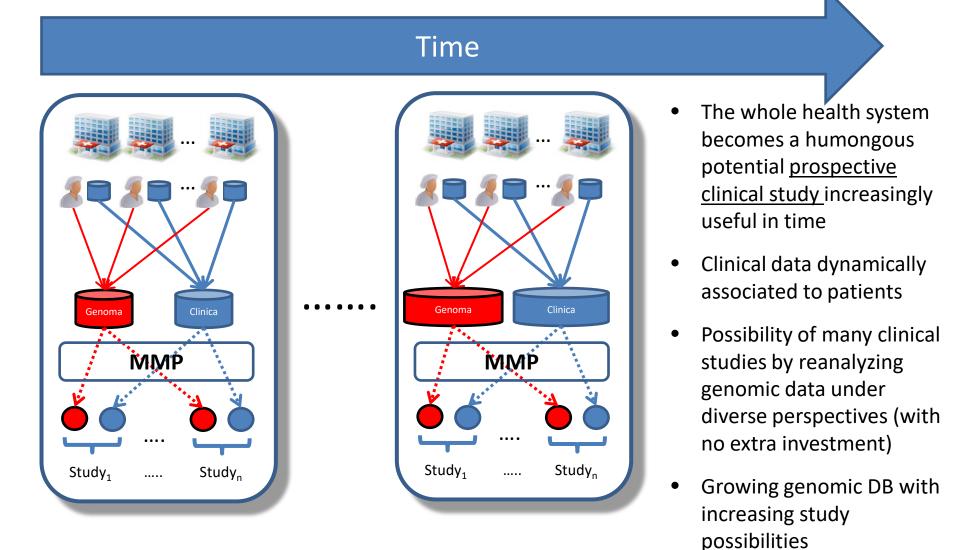
Personalized Medicine Model without universal eHR...



- Each study requires of a specific genomic and clinical data collection into an external database
- Static clinical data (e.g. if a control becomes a case the external DB will not be updated)
- Limited genomic data reuse for purposes different from the original study
- Model of GEL (100,000 genomes) Catalonian Genomic initiative, NaGen, etc.

... is actually only a clinical study

A personalized medicine model <u>must</u> integrate genomic data and universal eHR



Achievements of the first phase of the Personalized Medicine Plan of Andalucía

- <u>Personalized Medicine Module:</u> Intuitive software for semiautomatic diagnosis using massive sequencing data in a similar way to other clinical tests
- Used for the diagnostic of the 700 more prevalent RDs in Andalusia
- <u>High performance</u> engines for genomic data management and visualization in collaboration with the GEL (100,000 genomes project)
- Change in the current protocols to adopt massive sequencing instead conventional Sanger sequencing tests where possible
- Starting the phase II: Cancer

Clinical Bioinformatics Area Fundación Progreso y Salud, Sevilla, Spain, and...

...the INB-ELIXIR-ES, National Institute of Bioinformatics and the BiER (CIBERER Network of Centers for Research in Rare Diseases)



https://www.slideshare.net/xdopazo/

ciberer ine

Personalized Medicine





