유전체 기술을 이용한 심혈관 질환 진단방법

2019. 12. 14.

전성원 (Sungwon Jeon)

KOGIC, UNIST, South Korea

KOGIC
KOrean Genomics Industrialization and Commercialization Center
게놈산업기술센터

#JCR2019
#Healing Heart symposium

UNIST
울산과학기술원
ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

유전체 기술을 이용한 심혈관 질환 진단방법

2019. 12. 14.

전성원 (Sungwon Jeon)

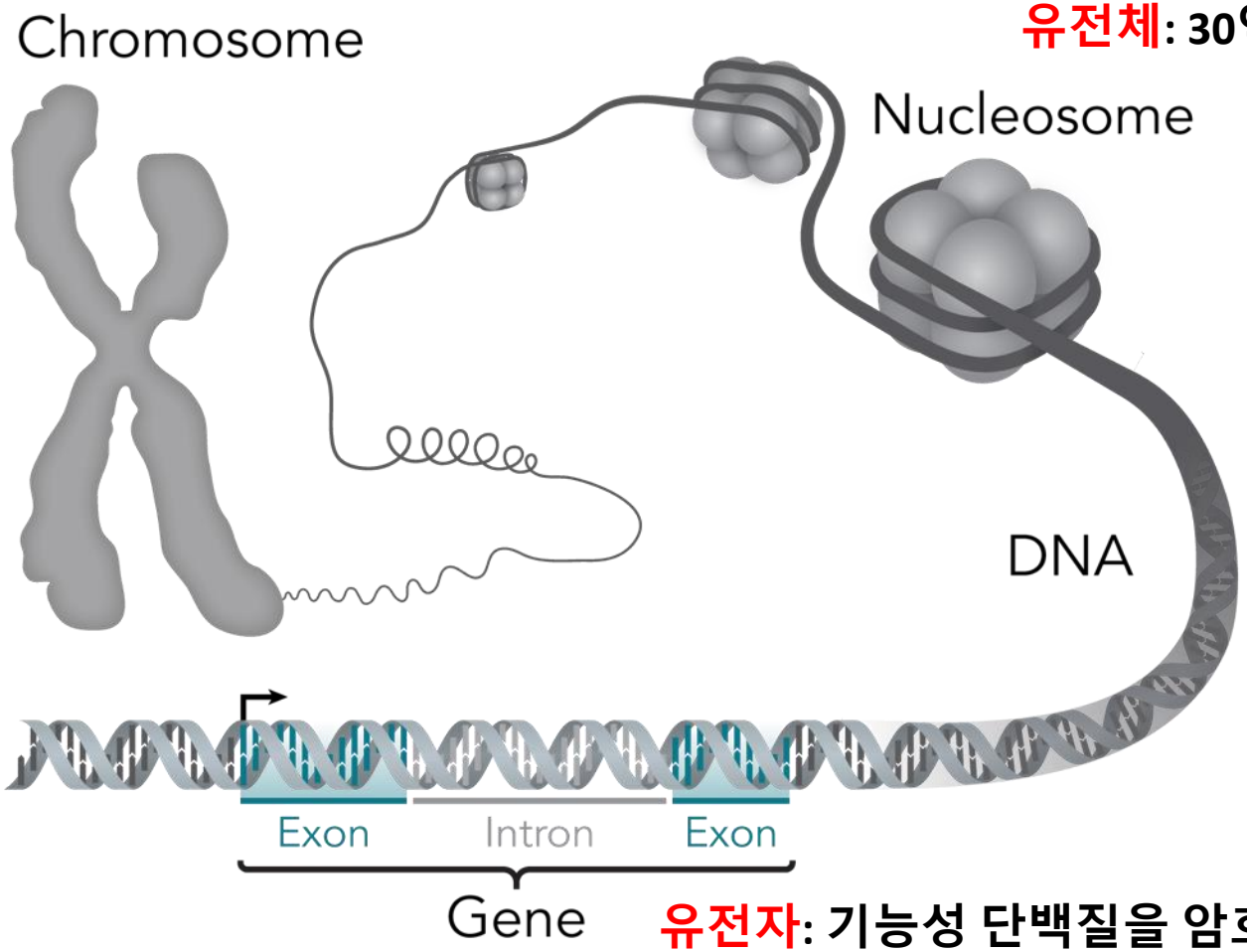
KOGIC, UNIST, South Korea

KOGIC
KOrean Genomics Industrialization and Commercialization Center
게놈산업기술센터

#JCR2019
#Healing Heart symposium

UNIST
울산과학기술원
ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

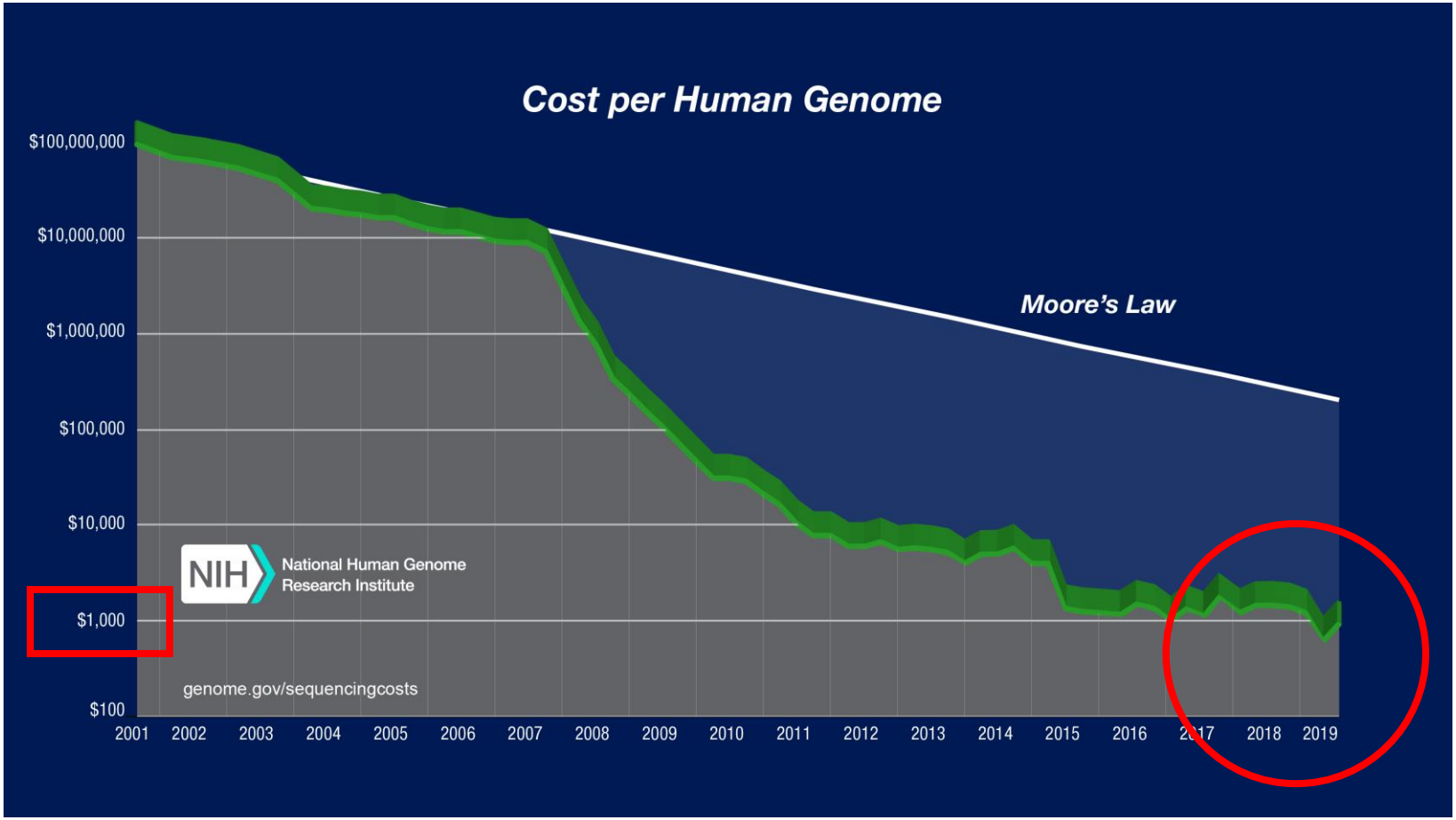
인간 게놈



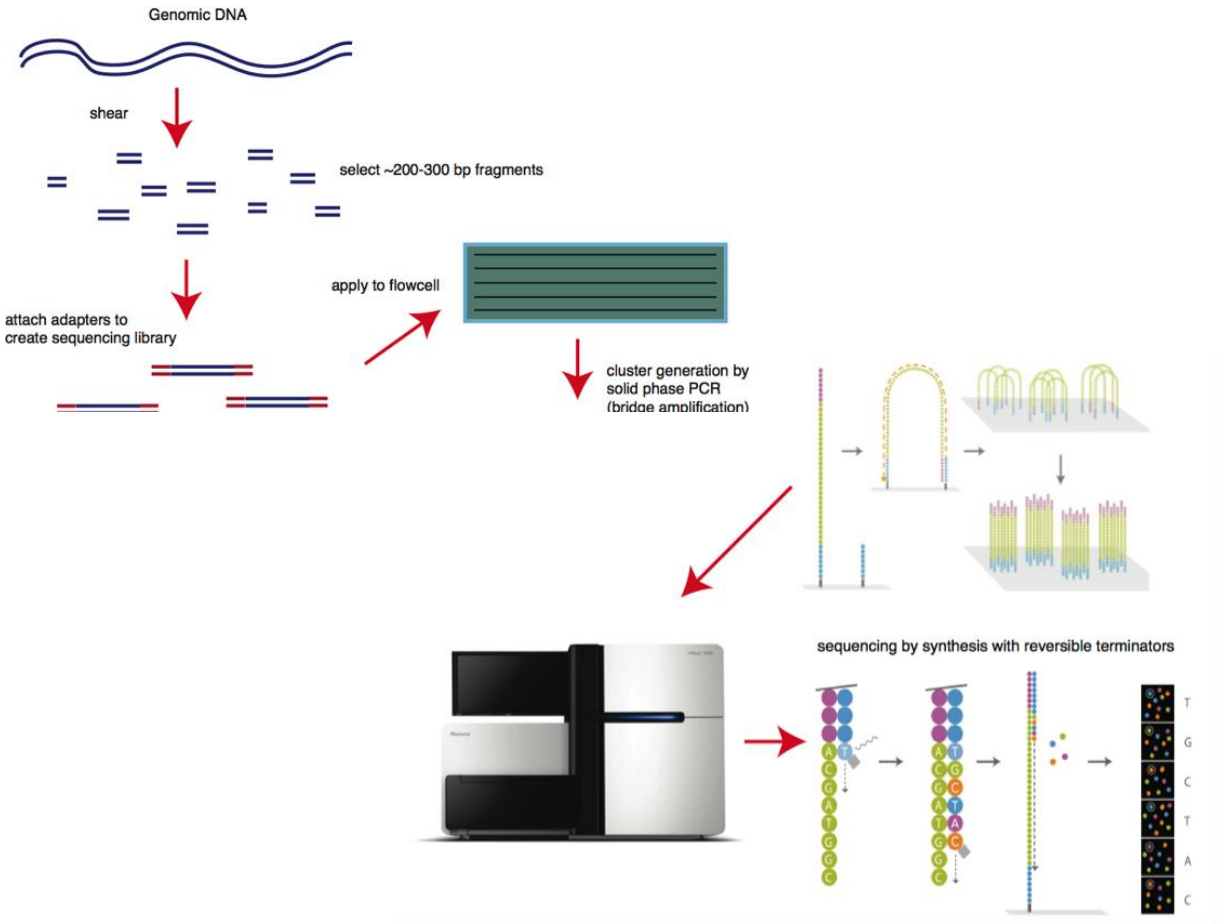
유전체: 30억 염기 서열 (사람경우)

유전자: 기능성 단백질을 암호화 하는 영역 (2만개)

Human whole genome sequencing cost

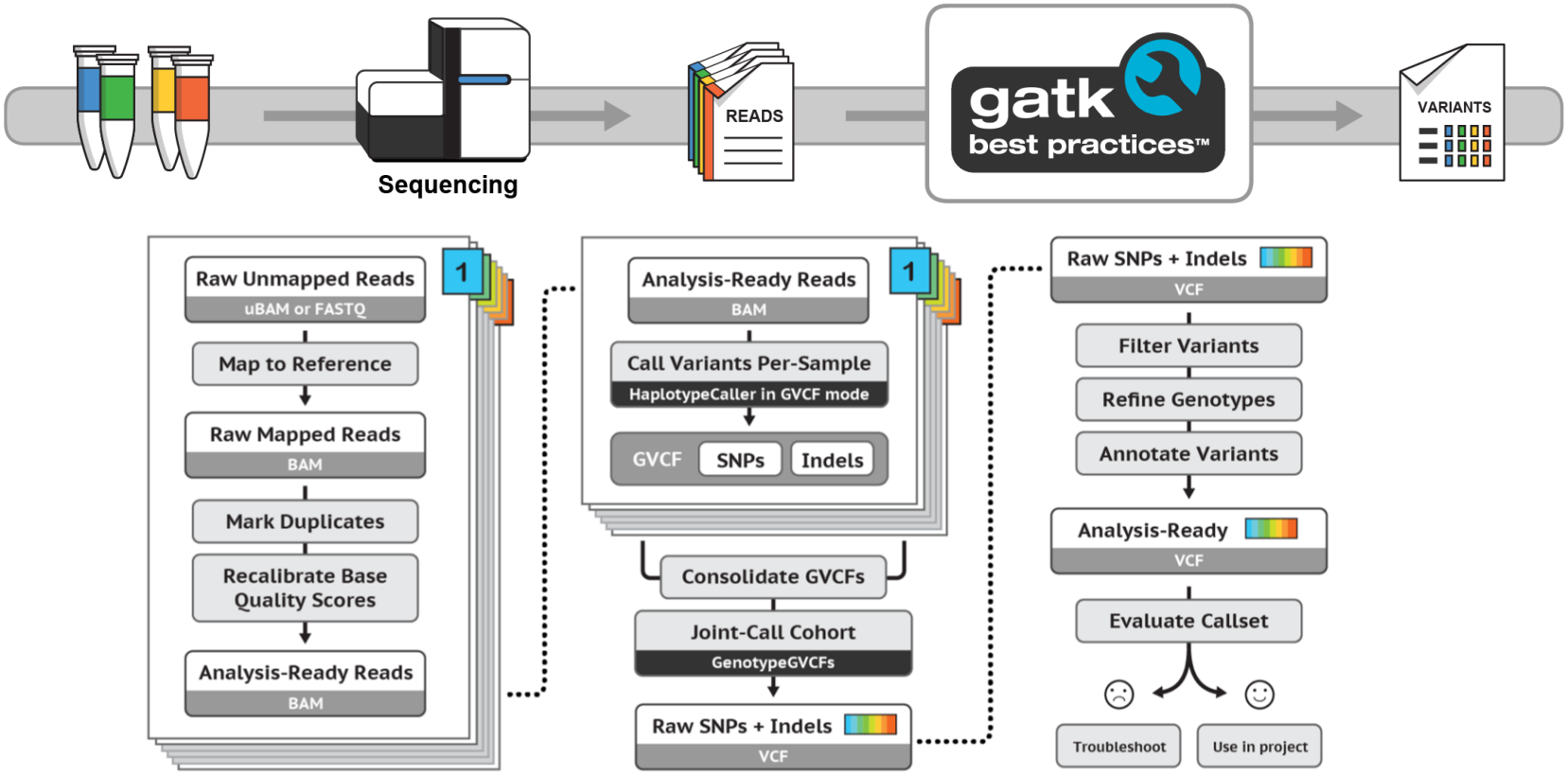


Next generation sequencing (NGS)



Bioinformatics (생정보학)

인간 게놈 분석 Pipeline



36 threads => 36 hours

https://software.broadinstitute.org/gatk/

https://www.intel.com/content/www/us/en/healthcare-it/solutions/documents/deploying-gatk-best-practices-paper.html

Human whole genome sequencing

Human genome sequencing data

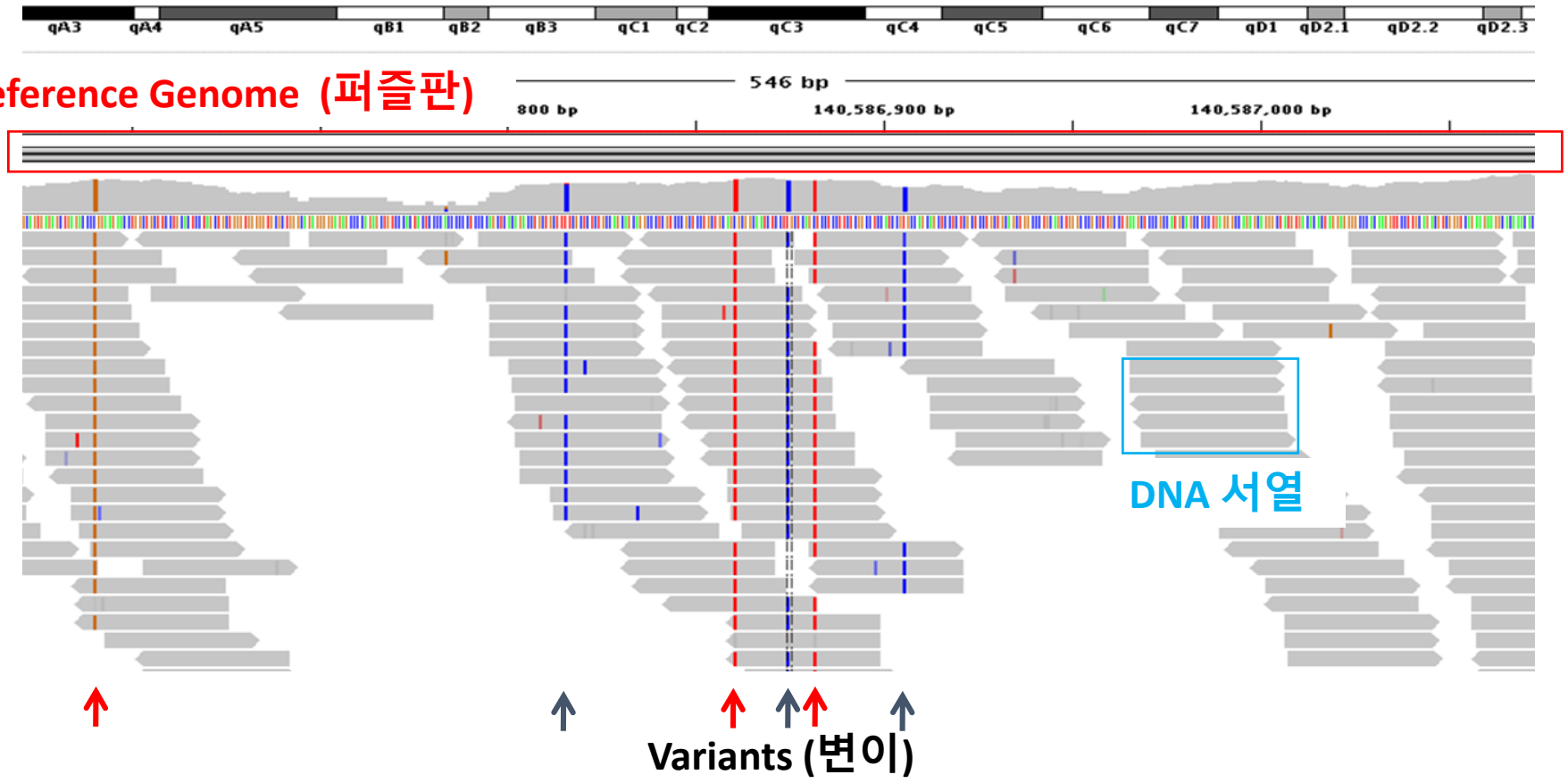
3Gb X 30배수 = 90Gb

생정보 기초분석 이후

0.5~1TB/sample

인간 게놈 분석

유전체: 30억 염기 서열 (사람경우)

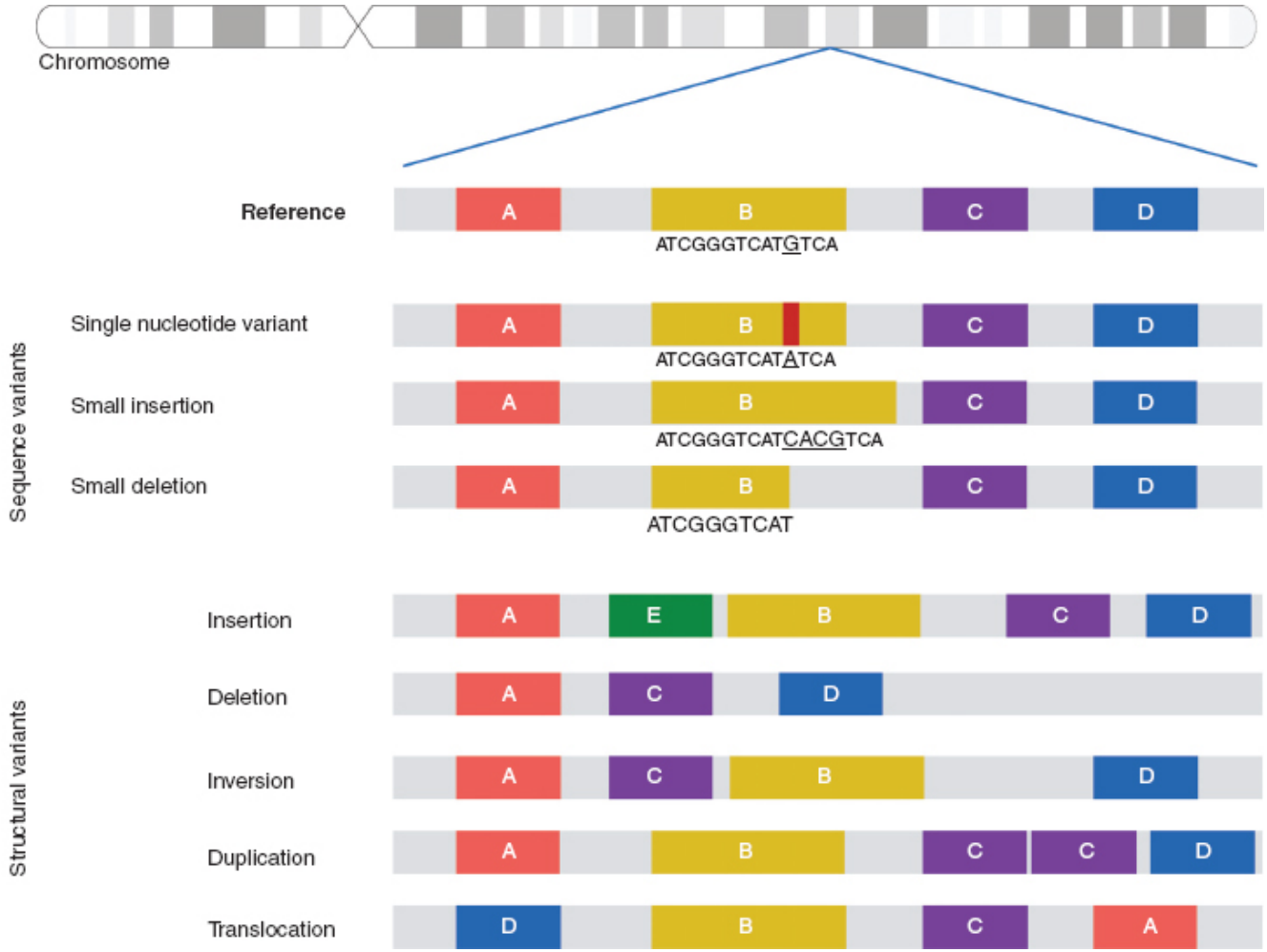


변이: 표준 서열과 다른 염기 서열
변이체: 총 변이의 합

보통 30X 정도는 해야한다.

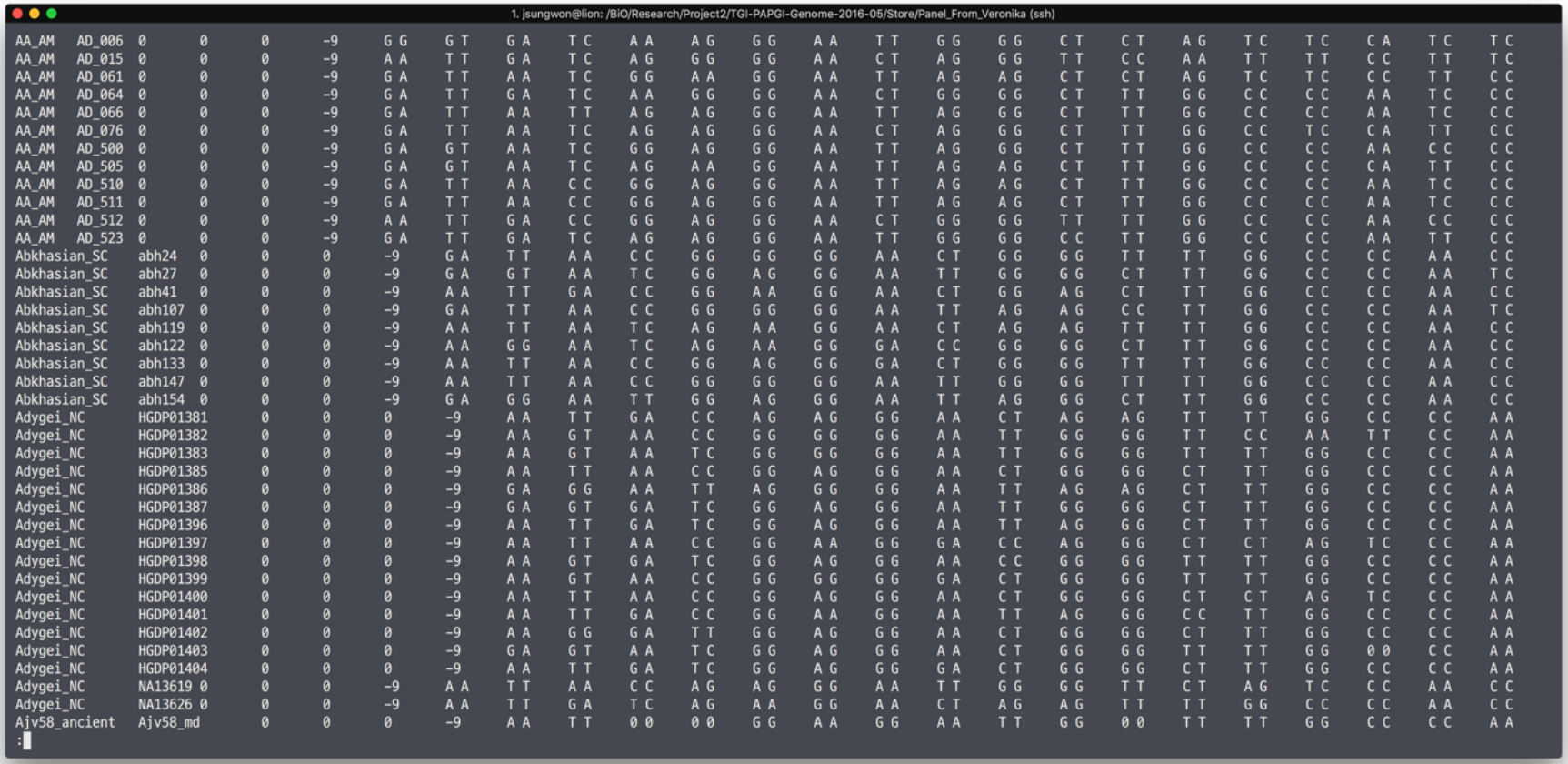
게놈 변이

https://neupsykey.com/genetics-3/



Variome (변이체)

샘플



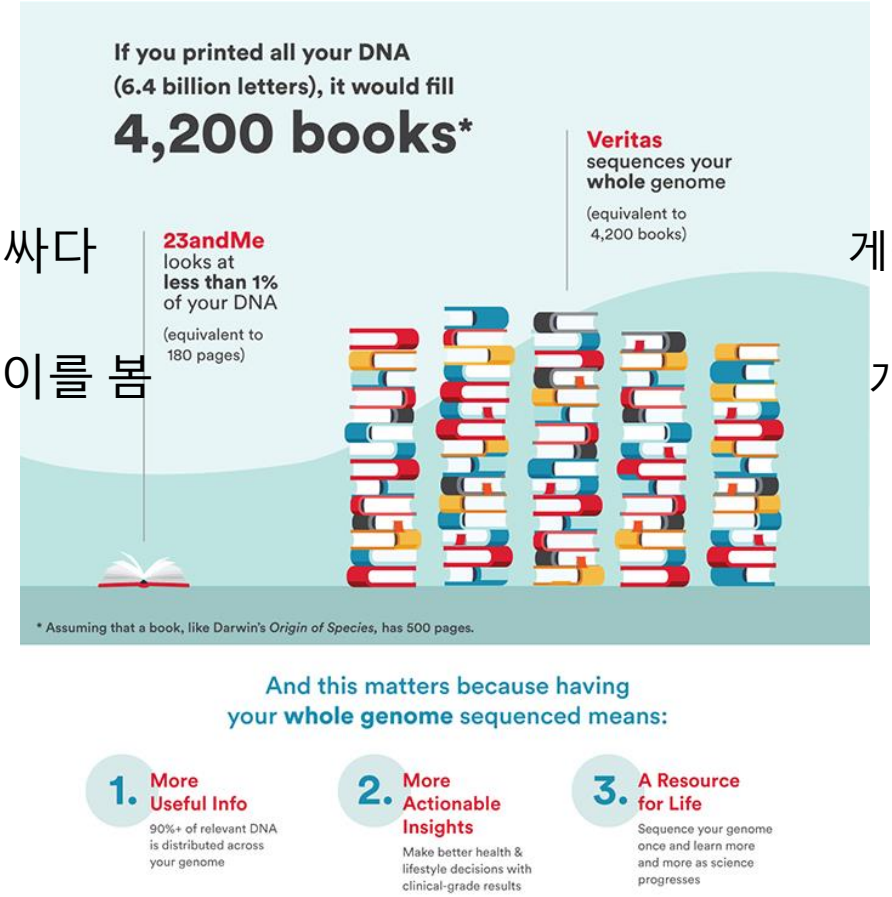
변이 정보

DNA Chip vs Whole genome sequencing (WGS)

DNA chip

가격이 상대적으로 싸다

대체로 Common 변이를 봄



WGS

게놈전체에 대한 변이 파악가능

개인 특이적인 변이도 발견가능

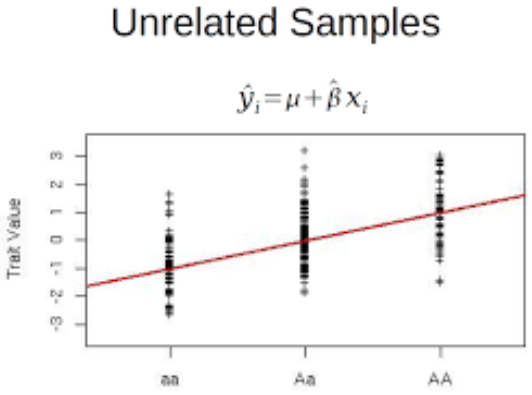
https://www.veritasgenetics.com/whats-difference-between-genetic-testing-technologies

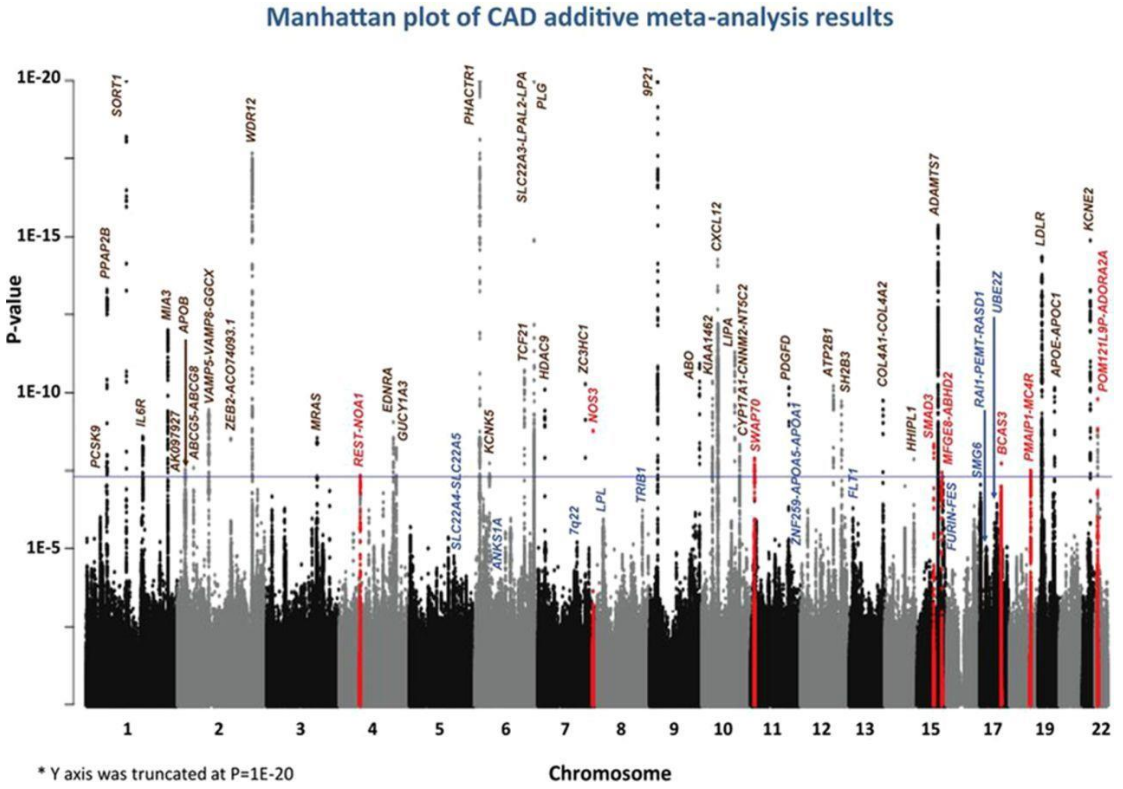
유전체기술로 심혈관질환 연구

예측?

진단?

Genome wide association study (GWAS)

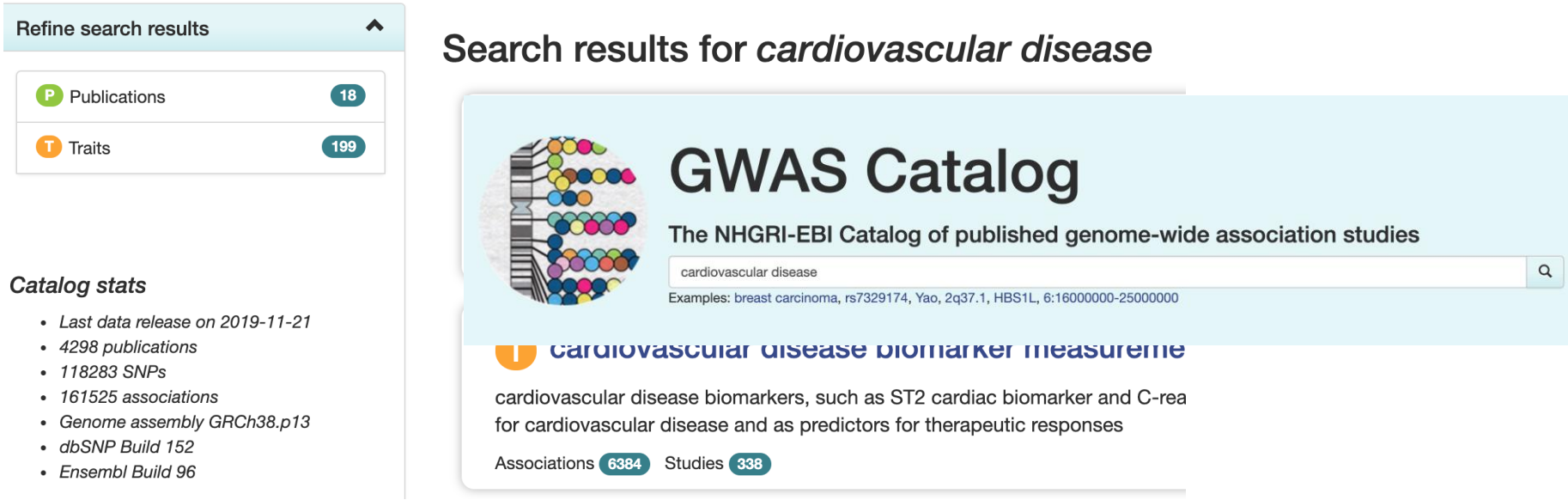




관련있는 변이 발견

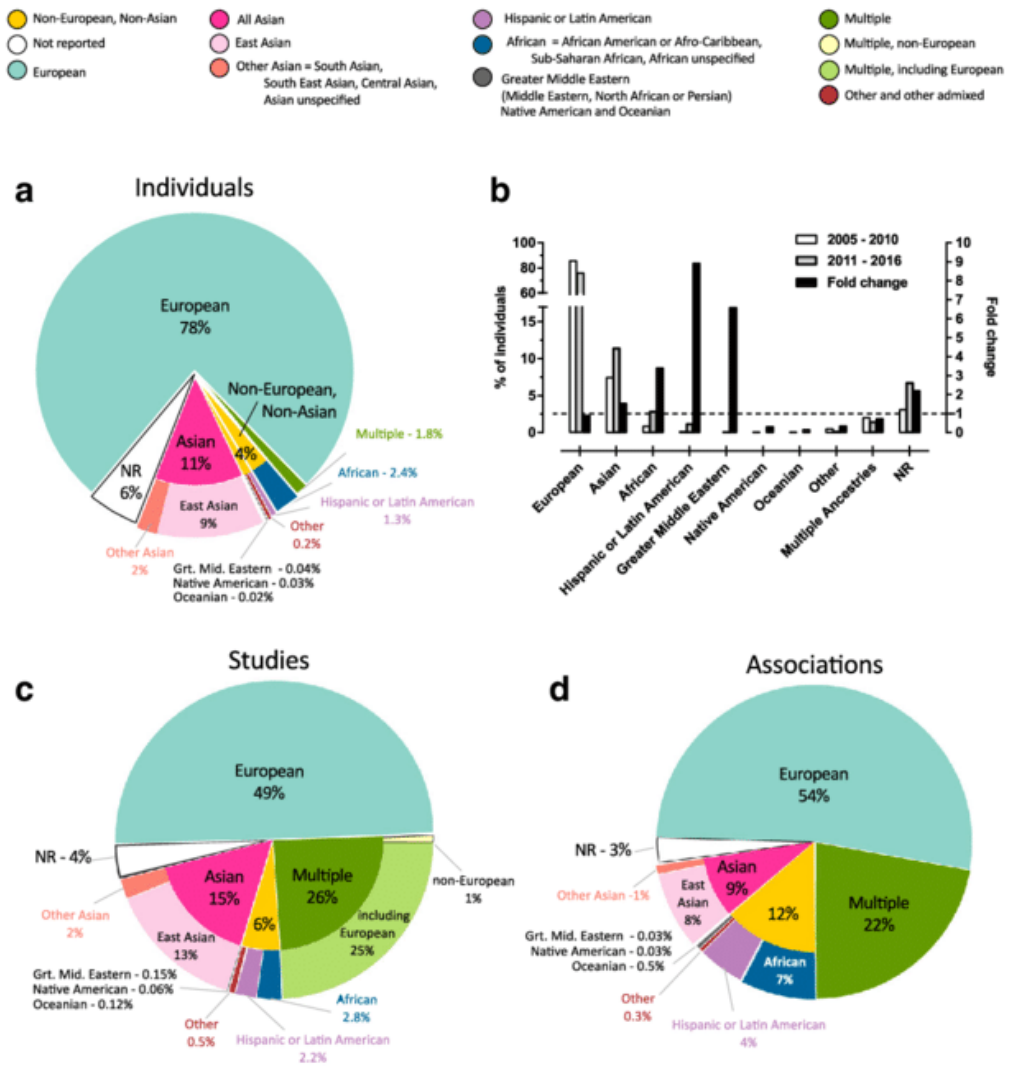
McPherson, Ruth, and Anne Tybjaerg-Hansen. "Genetics of coronary artery disease." *Circulation research* 118.4 (2016): 564-578.

CVD GWAS catalog variants



4,298 publications, 118,282 SNPs from GWAS catalog

Skewed population distribution



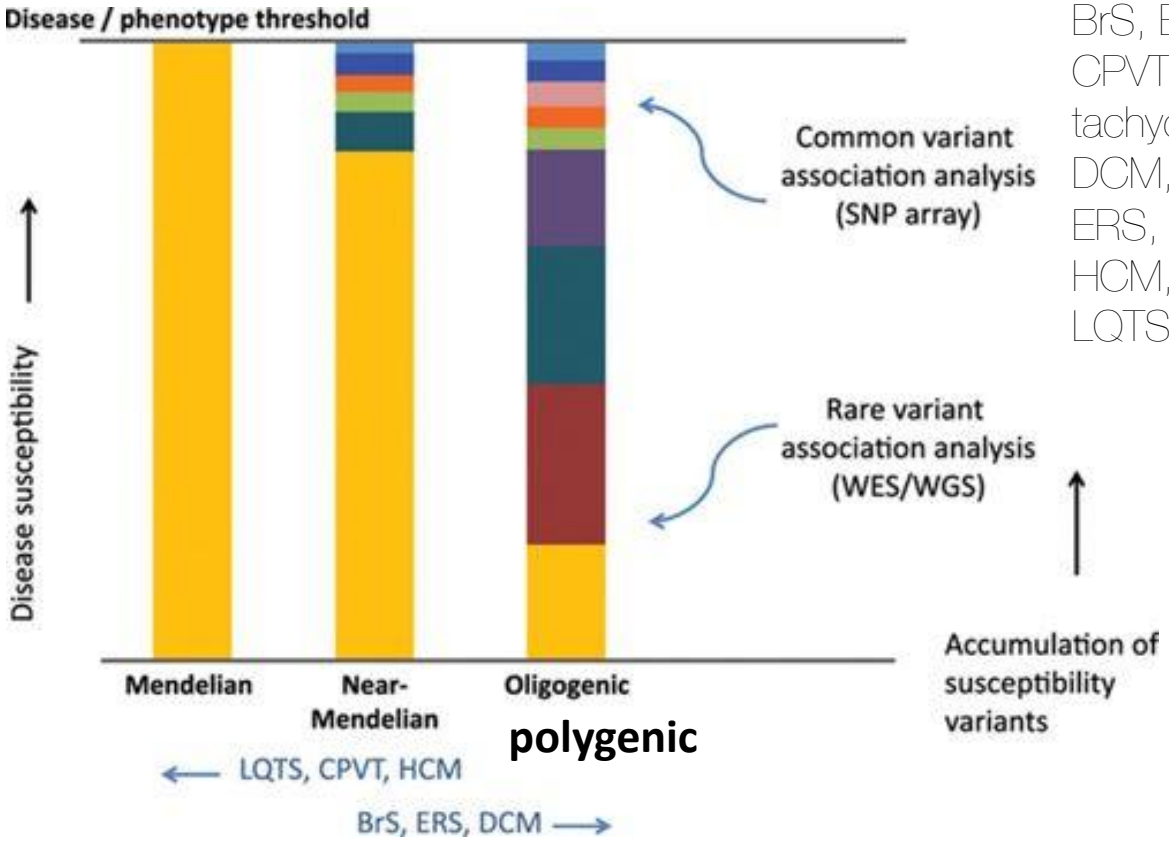
한국인, 아시아인 연구의 부족

KOGIC
KOrean Genomics Industrialization and Commercialization Center
게놈산업기술센터

Morales, Joannella, et al. "A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog." *Genome biology* 19.1 (2018): 21.

Monogenic vs polygenic

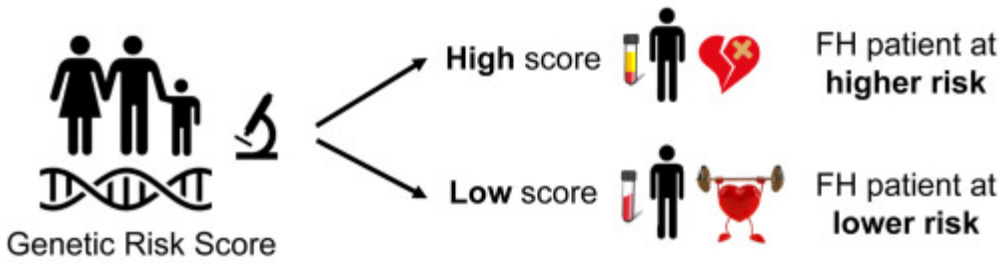
Bezzina, Connie R., Najim Lahrouchi, and Silvia G. Priori. "Genetics of sudden cardiac death." *Circulation research* 116.12 (2015): 1919-1936.



BrS, Brugada syndrome
CPVT, catecholaminergic polymorphic ventric
tachycardia
DCM, dilated cardiomyopathy
ERS, early repolarization syndrome
HCM, hypertrophic cardiomyopathy
LQTS, long-QT syndrome

Polygenic (genetic) risk score

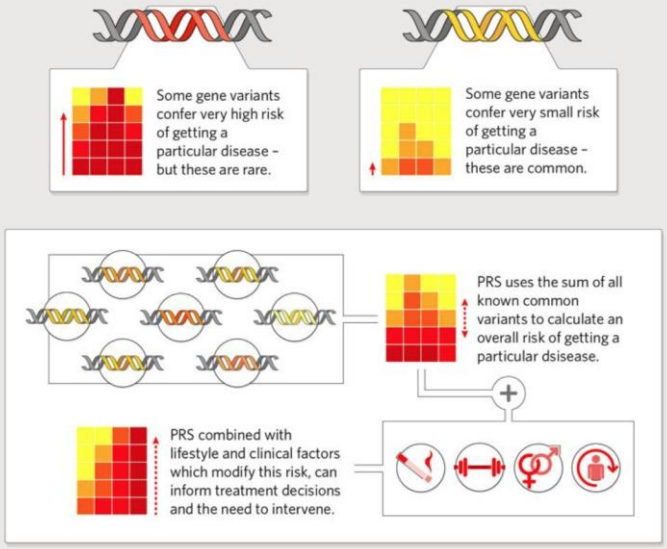
$$\hat{S} = \sum_{j=1}^{m} X_j \hat{\beta}_j,$$



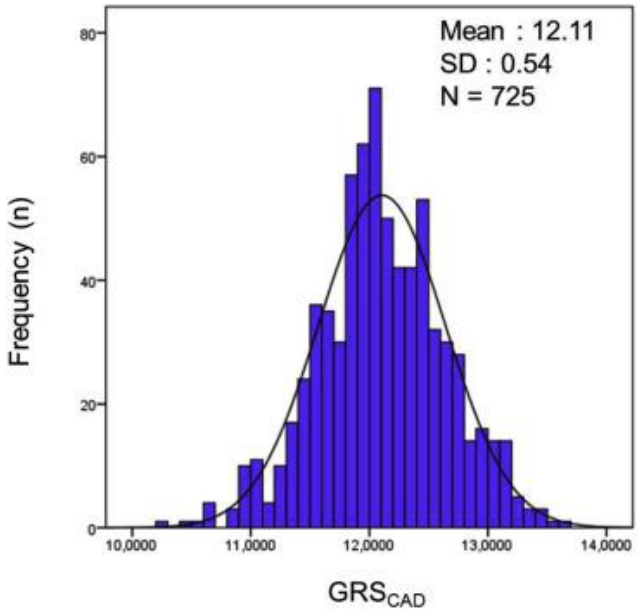
CLINICAL APPLICATION OF PRS

A polygenic risk score (PRS) is calculated from many small genetic variants, and can often be modified by lifestyle factors.

https://www.nature.com/articles/d42473-019-00270-w



Polygenic factors can influence the FH phenotype



Paquette, Martine, et al. "Polygenic risk score predicts prevalence of cardiovascular disease in patients with familial hypercholesterolemia." *Journal of clinical lipidology* 11.3 (2017): 725-732.

AI for CVD Risk prediction

RESEARCH ARTICLE

Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants

Ahmed M. Alaa[1]*, Thomas Bolton[2,3], Emanuele Di Angelantonio[2,3], James H. F. Rudd[4], Mihaela van der Schaar[1,5,6]

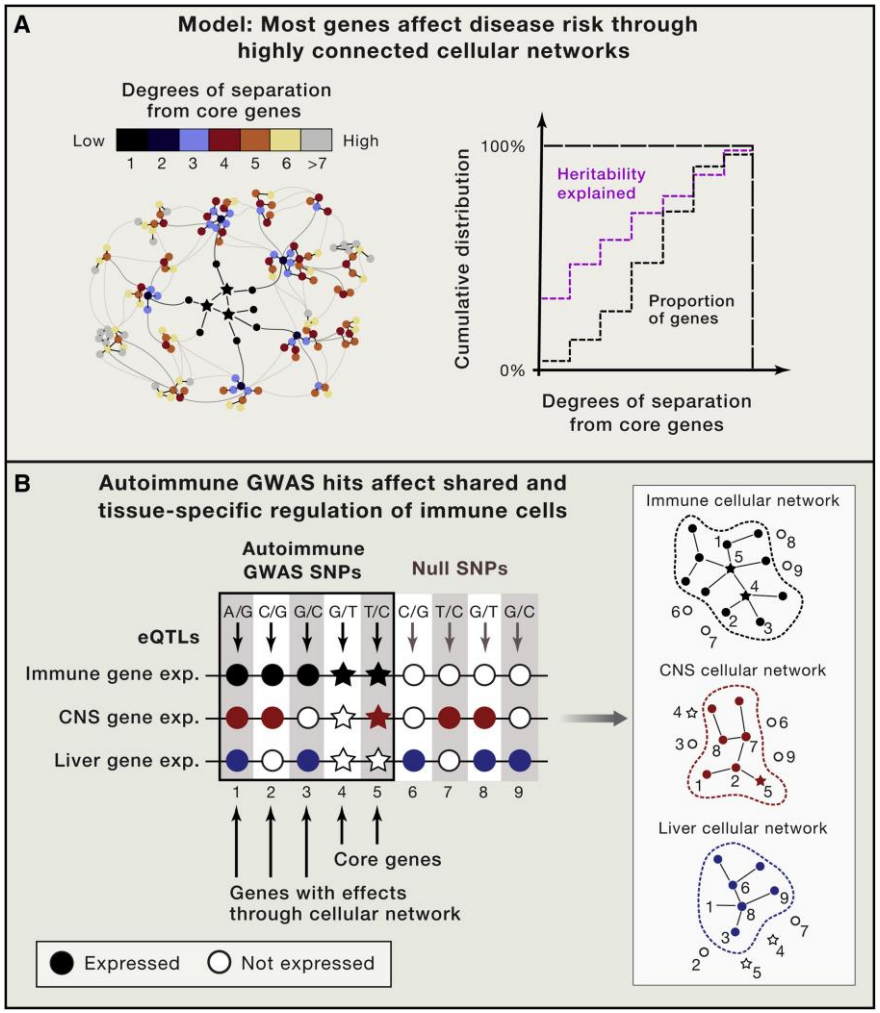
1 University of California Los Angeles, Los Angeles, California, United States of America, **2** Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom, **3** National Institute for Health Research (NIHR) Blood and Transplant Research Unit (BTRU) in Donor Health and Genomics, University of Cambridge, Cambridge, United Kingdom, **4** Department of Cardiovascular Medicine, University of Cambridge and Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom, **5** University of Oxford, Oxford, United Kingdom, **6** Alan Turing Institute, London, United Kingdom

* ahmedmalaa@ucla.edu

Model	AUC-ROC	Absolute AUC-ROC Change
Framingham Score	0.724 ± 0.004	Baseline model
Cox PH Model (7 core variables)	0.734 ± 0.005	+ 1.0%
Cox PH Model (all variables)	0.758 ± 0.005	+ 3.4%
Support Vector Machines	0.709 ± 0.061	- 1.5%
Random Forest	0.730 ± 0.004	+ 0.6%
Neural Networks	0.755 ± 0.005	+ 3.1%
AdaBoost	0.759 ± 0.004	+ 3.5%
Gradient Boosting	0.769 ± 0.005	+ 4.5%
AutoPrognosis (7 core variables)	0.744 ± 0.005	+ 2.0%
AutoPrognosis (369 non-lab. variables)	0.761 ± 0.005	+ 3.7%
AutoPrognosis (104 lab. variables)	0.735 ± 0.008	+ 1.1%
AutoPrognosis (all variables)	0.774 ± 0.005	+ 5.0%

Omnigenic model

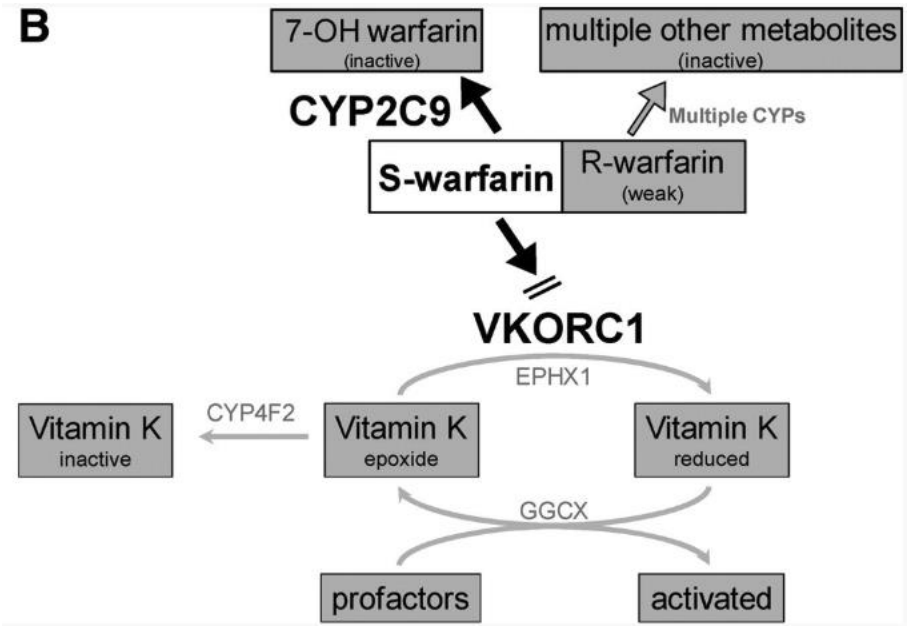
Boyle, Evan A., Yang I. Li, and Jonathan K. Pritchard. "An expanded view of complex traits: from polygenic to omnigenic." *Cell* 169.7 (2017): 1177-1186.



We propose that gene regulatory networks are sufficiently interconnected such that all genes expressed in disease-relevant cells are liable to affect the functions of core disease-related genes and that most heritability can be explained by effects on genes outside core pathways.

무한히 많은 작은 효과들

Pharmacogenomics : Warfarin

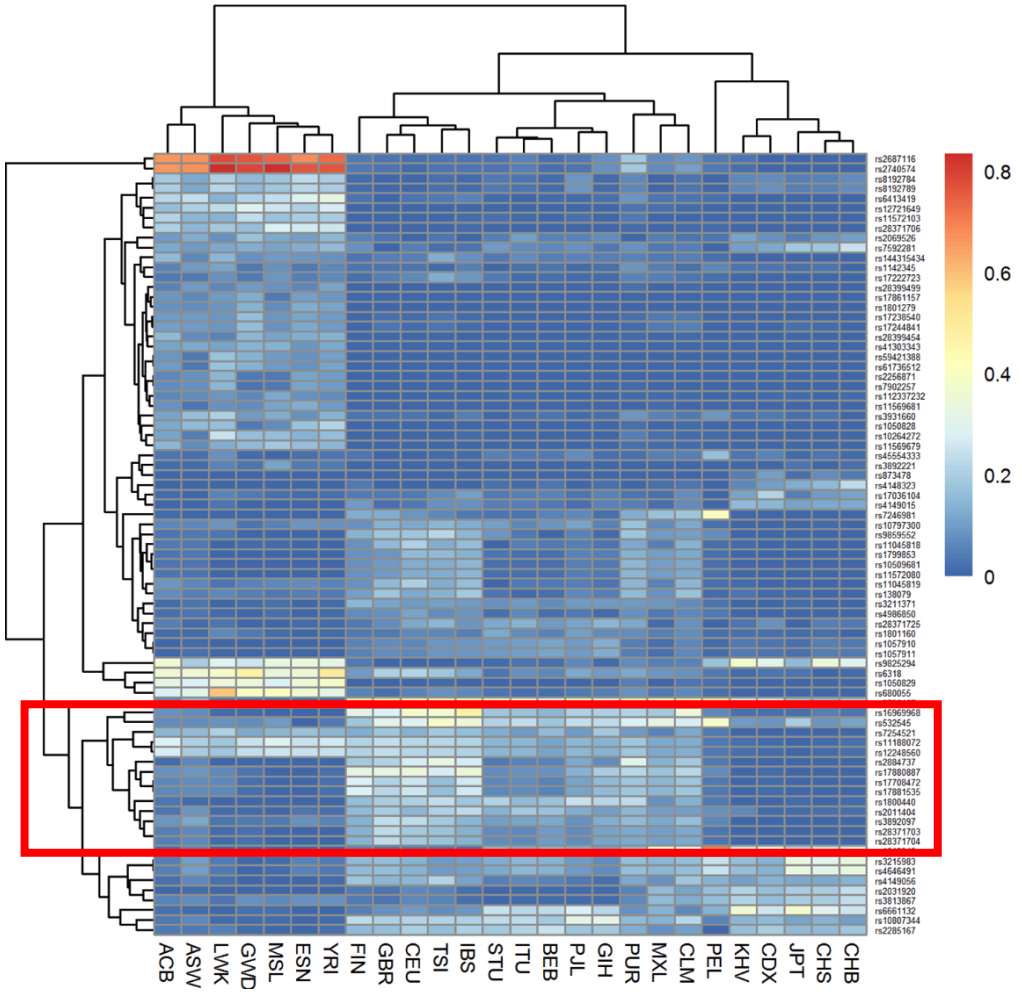


variants in CYP2C9 and VKORC1 contribute to variable warfarin dosage.

Roden, Dan M., et al. "Cardiovascular pharmacogenomics." *Circulation research* 109.7 (2011): 807-820.

Pharmacogenomics

인종마다 변이의
빈도가 차이남



아프리카 유럽 아메리카 동아시아

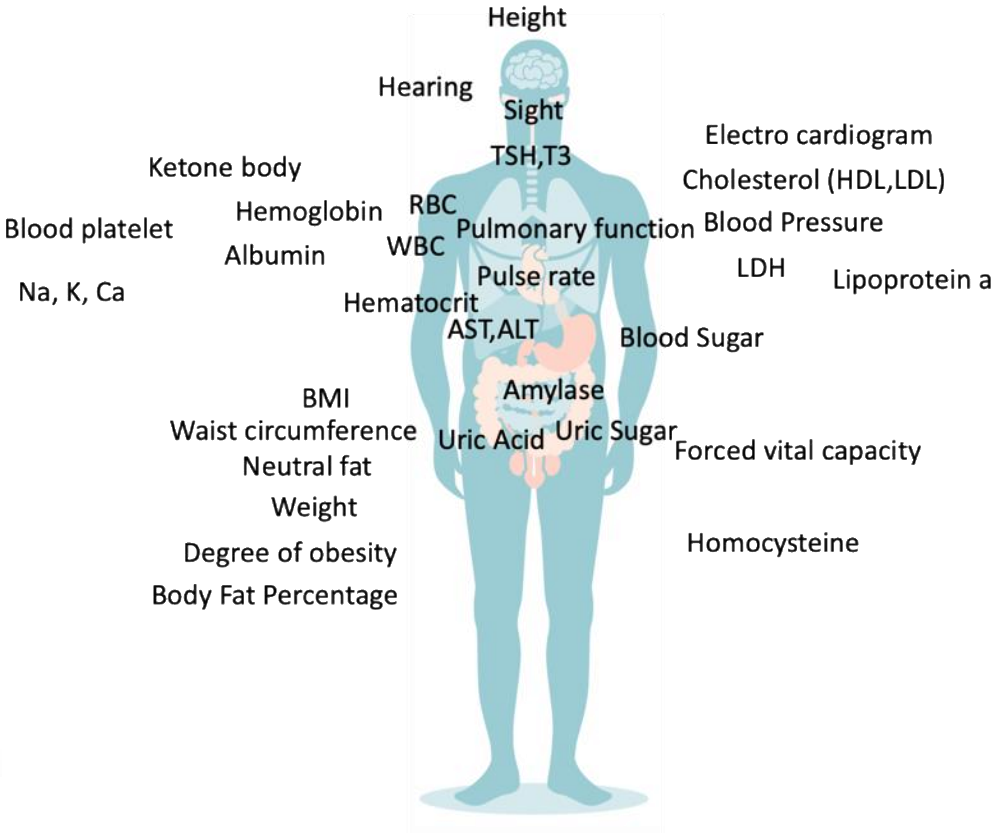
Minor Allele Frequency heatmap for 79 Pharmacogenomics variants that are common (MAF> = 0.1) and rare (MAF< = 0.005) in at least one 1KG population.

Lakiotaki, Kleanthi, et al. "Exploring public genomics data for population pharmacogenomics." *PloS one* 12.8 (2017): e0182138.

Korean genome database

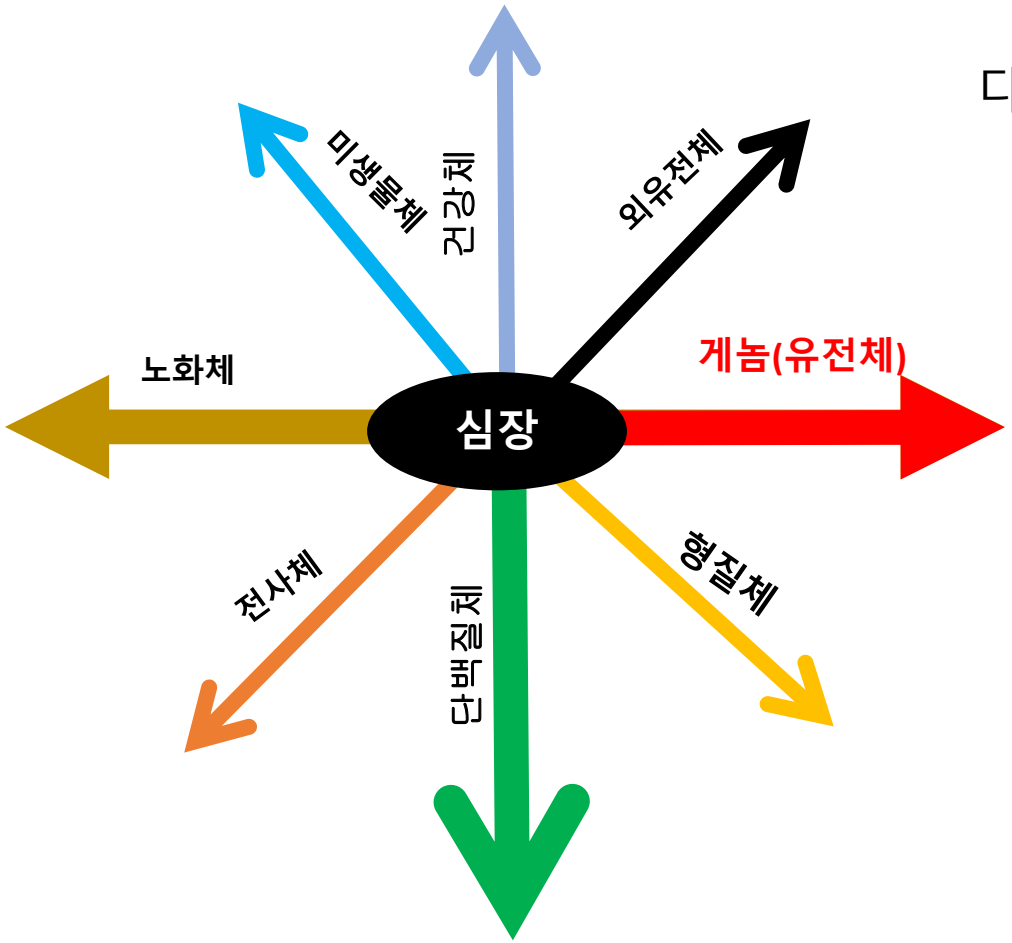
울산 만명게놈 프로젝트

국민 게놈



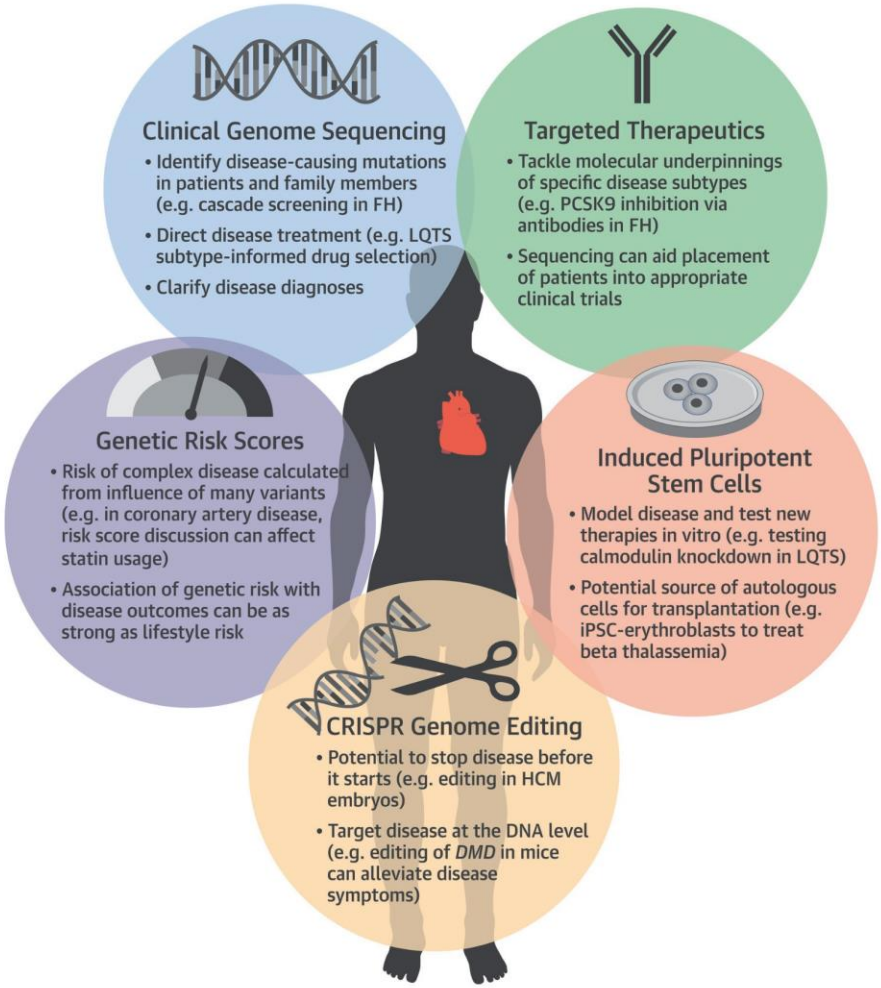
건강 검진 정보

Multi-omics



다중 오믹스를 통한 심장게놈 연구

Precision/Personalized medicine



개인 게놈 정보를
활용한
정밀/맞춤의학

Dainis, Alexandra M., and Euan A. Ashley. "Cardiovascular precision medicine in the genomics era." *JACC: Basic to Translational Science* 3.2 (2018): 313-326.