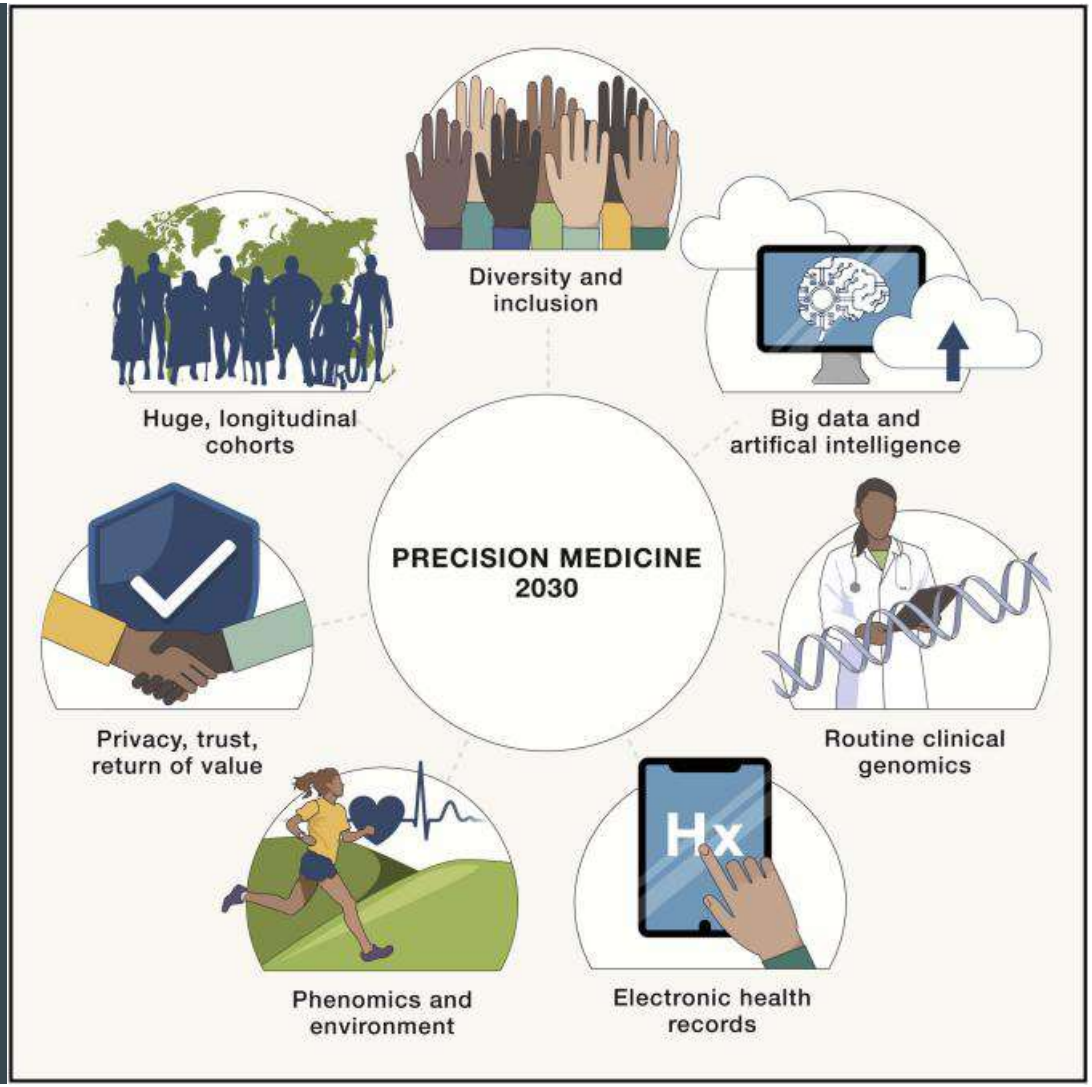
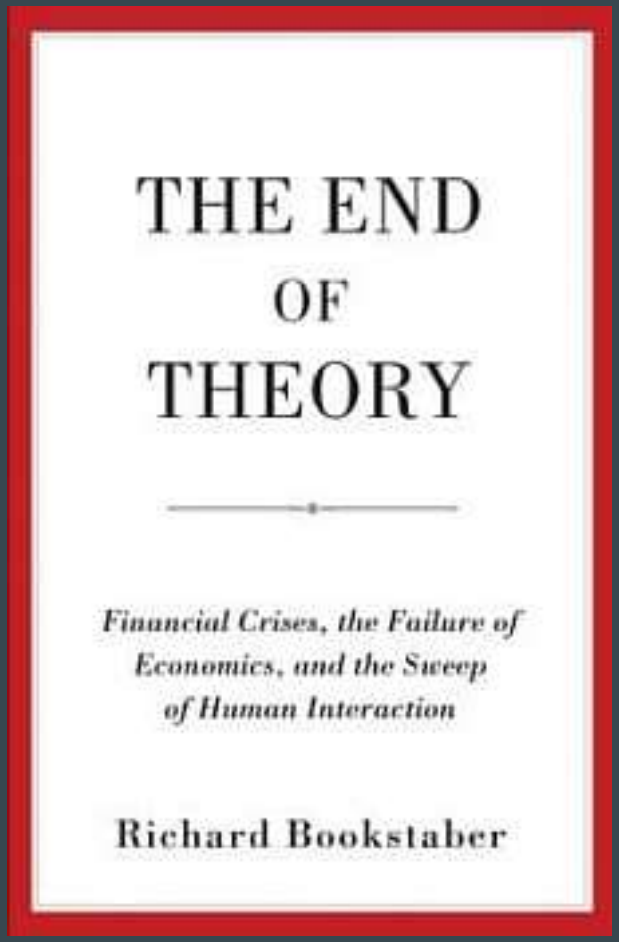
臺灣人體生物資料庫

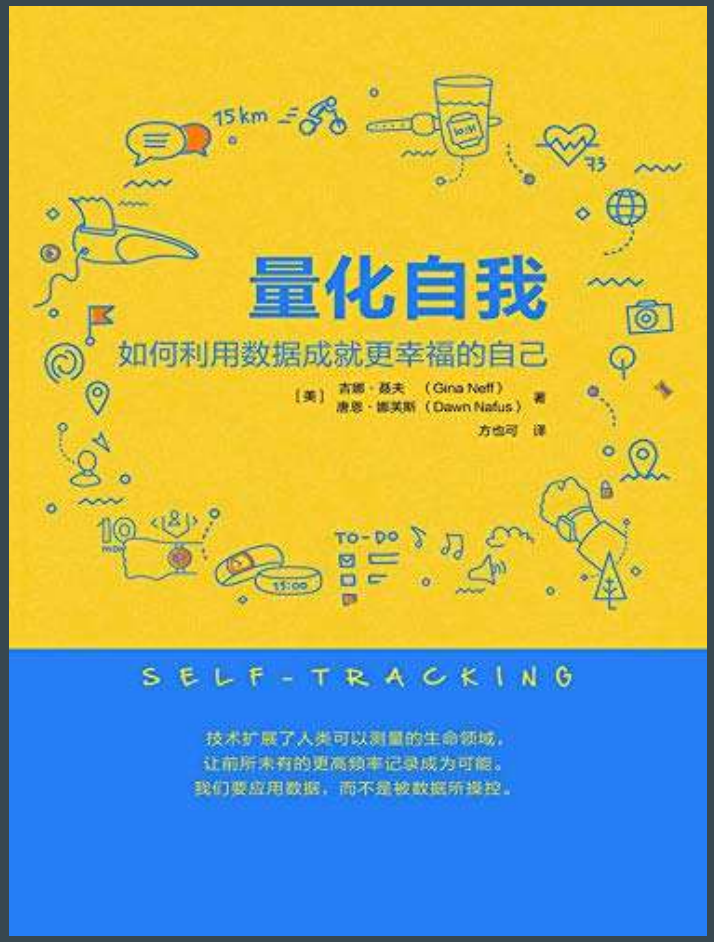
蘇明威
wei@ibms.sinica.edu.tw

Precision medicine in 2030—seven ways to transform healthcare

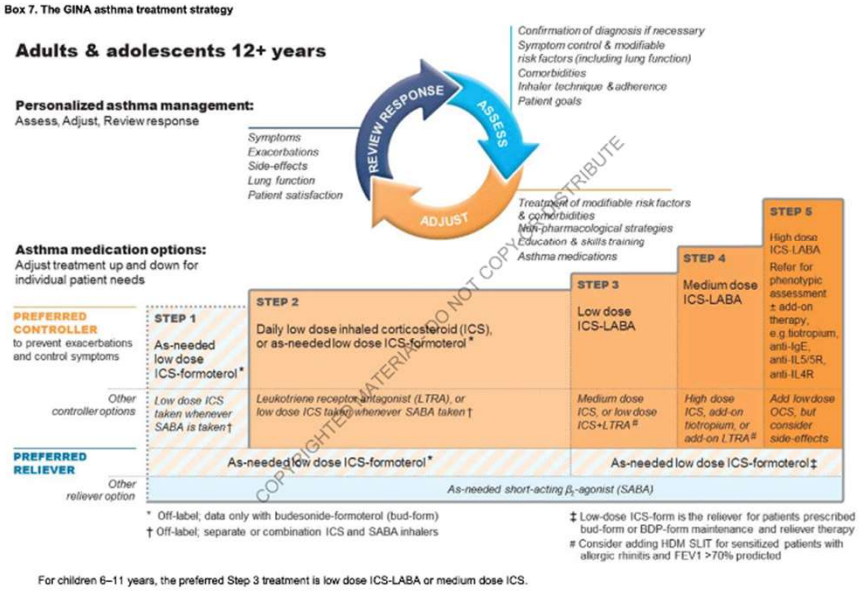


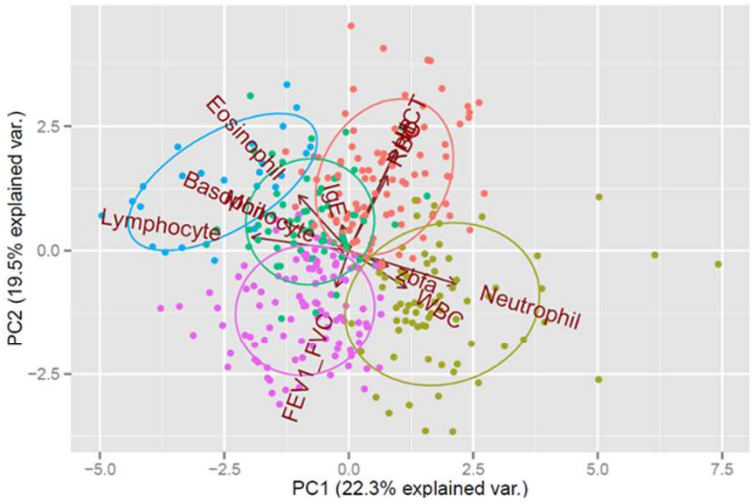
Transparency



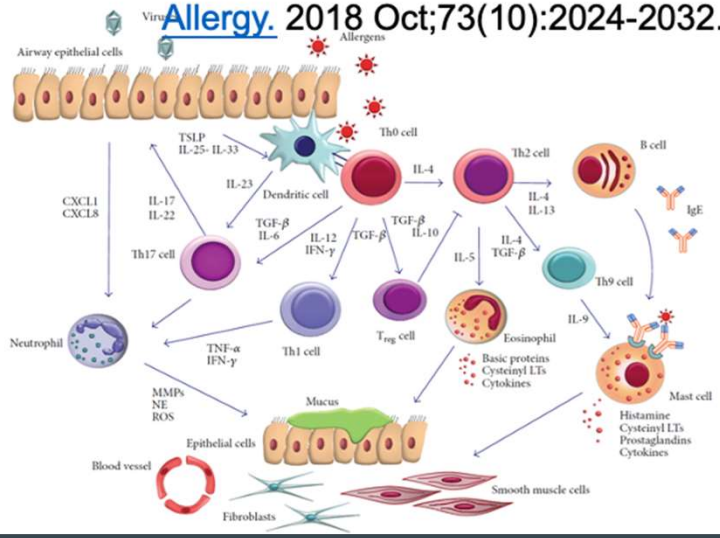


Precision medicine

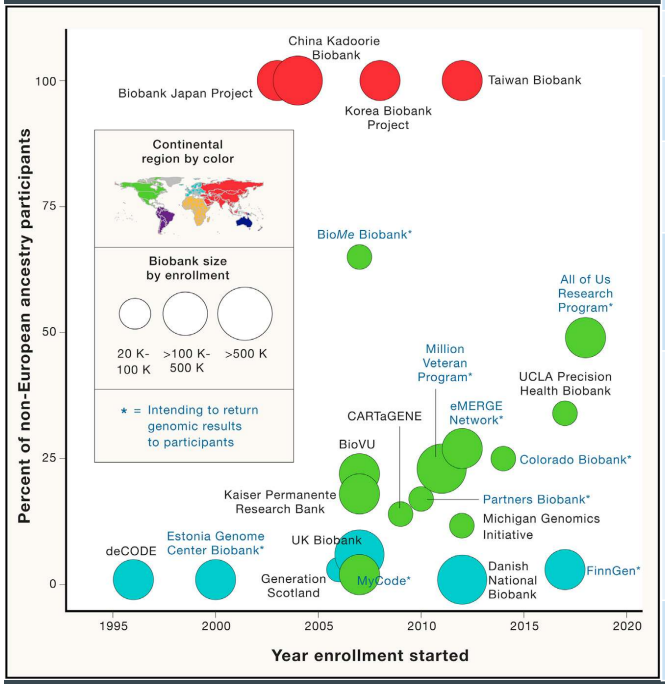




Allergy. 2018 Oct;73(10):2024-2032.



Nature 520, 609–611



	Taiwan Biobank (TWB)	Taiwan Precision Medicine Initiative (TPMI)	NHRI-資料庫整合平台 (資料來源為109年綱要計畫網站中可查詢到的摘要)
目標族群	20萬社區民眾 (>20 歲)	12家醫院110萬就醫或健診者	整合台灣現有之人體生物資料庫, 建立一個國家級人體生物資料庫整合平台
檢體收集與保存	血漿、尿液、DNA、及血球細胞等	利用醫院剩餘血液檢體抽取DNA, 不保留檢體	將各資料庫檢體收集流程以及檢體品質達成一致性的標準
資料收集	問卷, 血液檢測數據, 追蹤時加做進階檢驗及影像檢驗	病歷資訊	建立充足且一致性的臨床資料
檢體加值	全基因體SNP定型及定序, 全基因體甲基化, HLA 定序, 代謝體, 塑化劑與三聚氰胺等環境因子	全基因體SNP定型	將以大量經費投注於人體生物資料庫檢體之加值服務
資料/檢體釋出與資訊平台	各項檢體/問卷/檢測數據/檢體加值資料供申請使用，可串聯健保資料庫	TPMI 建立 Data-Lake, 醫院將病歷資訊匯入, 資料去識別, 只進不出, 提供醫院評估病患醫療策略及學者研究之用.	建立一個龐大完整的生醫大數據，符合生技製藥，人工智慧，輔助醫療等產業界的需求。
預期效益	1. 長期追蹤、以族群為基礎(population-based)的大型研究cohort。 2. 以UK-Biobank 為標竿，提高國際能見度及競爭力，期許成為最具規模的華人生物資料庫。 3. 資料庫的巨量資料，預期可激發許多創新性的技術開發及學術研究。	臨床應用層面 直接應用於臨床實務，以助增加用藥精確性並篩檢特定疾病。 學術研究層面 收集台灣華人專屬的數據，以找尋國人常見疾病的風險因子及藥物不良反應相關因子，並開發專屬台灣華人的基因型鑑定晶片。	提供學術界和產業界透過此平台申請使用，藉由人工智慧，有利於各種新藥開發以及建立輔助醫療之應用程式，將擁有龐大商機。
法遵	人體生物資料庫管理條例	人體研究法	?個資法

Cell. 2019 Mar 21; 177(1): 58–69.

長期追蹤, 亞洲唯一可對外提供資料及檢體

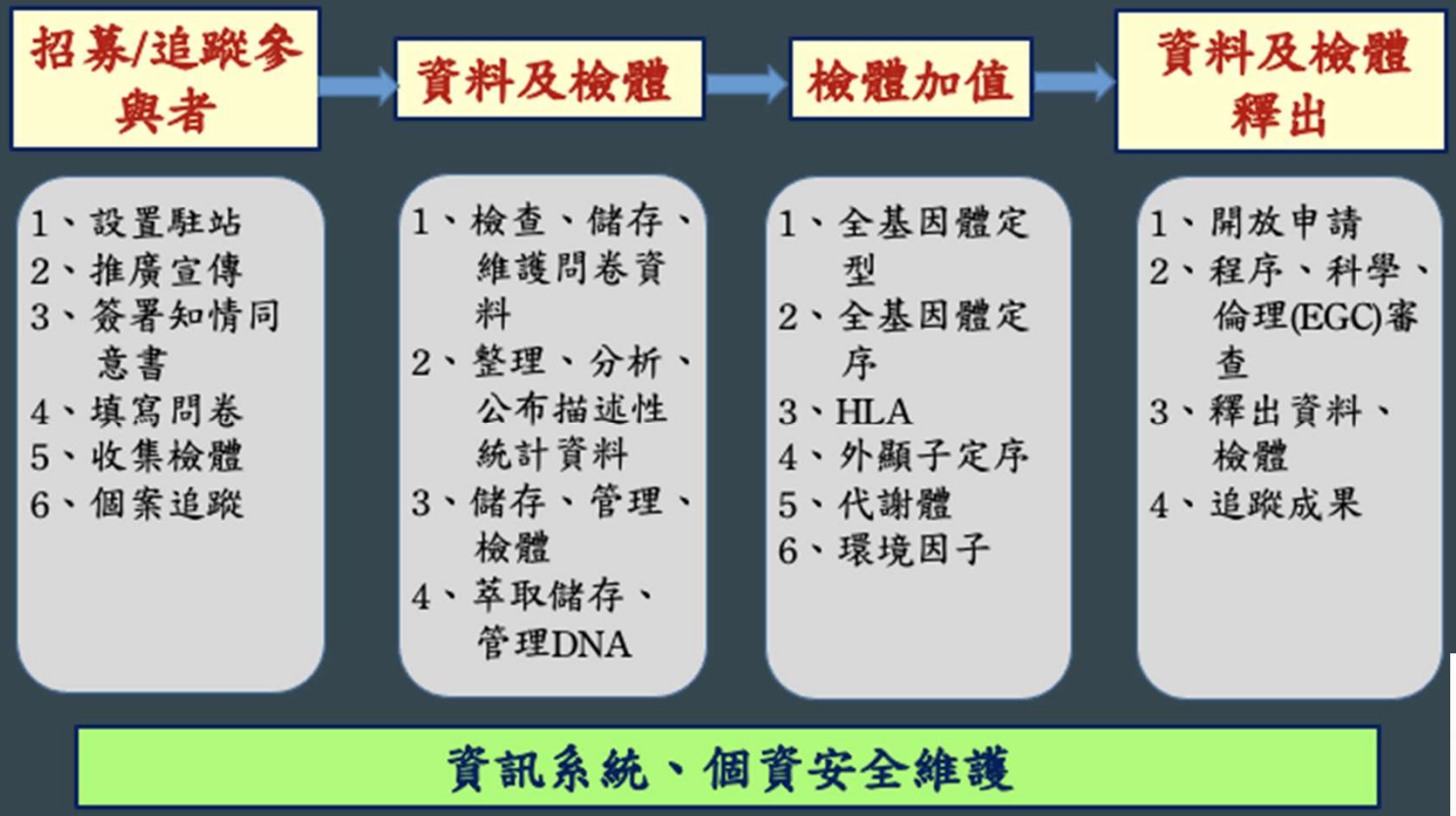
Why Taiwan ?

- Holo Taiwanese
- Hakka Taiwanese
- Mainianders
- Taiwanese indigenous proples
- Taiwanese new immigrants

中華民國
總統府 Office of the President
Republic of China (Taiwan)
www.president.gov.tw

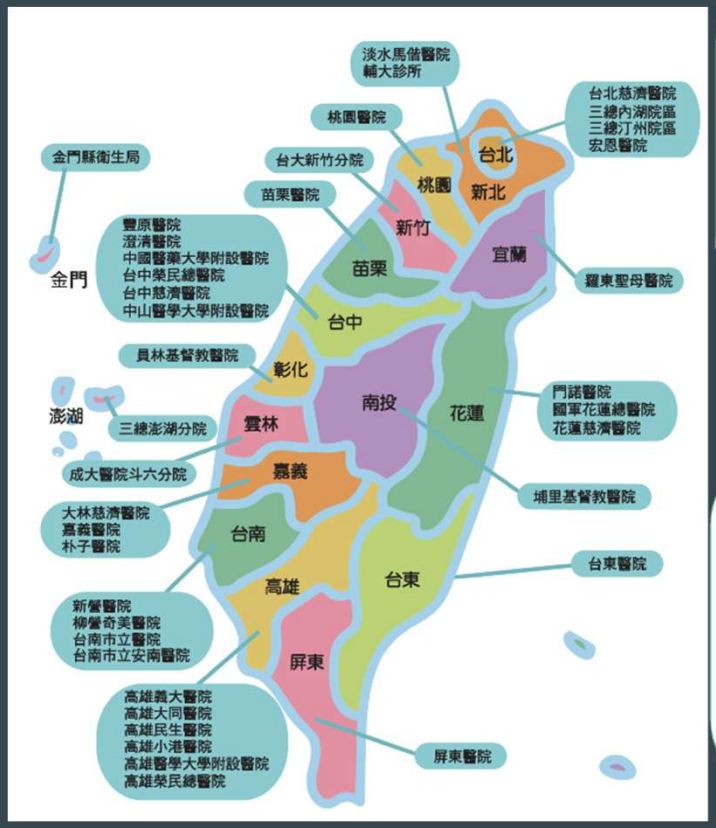
https://ogme.edu.tw/lc/culturalGroups
https://tendashsix.com/taiwan-medical-service-ranked-first/

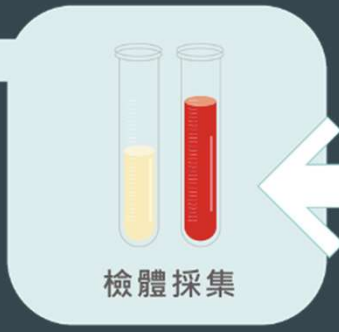
TWB infrastructure



臺灣 Taiwan Biobank

參與流程 (https://reurl.cc/n5j5Vv)





data and tubes

~1.5 PB

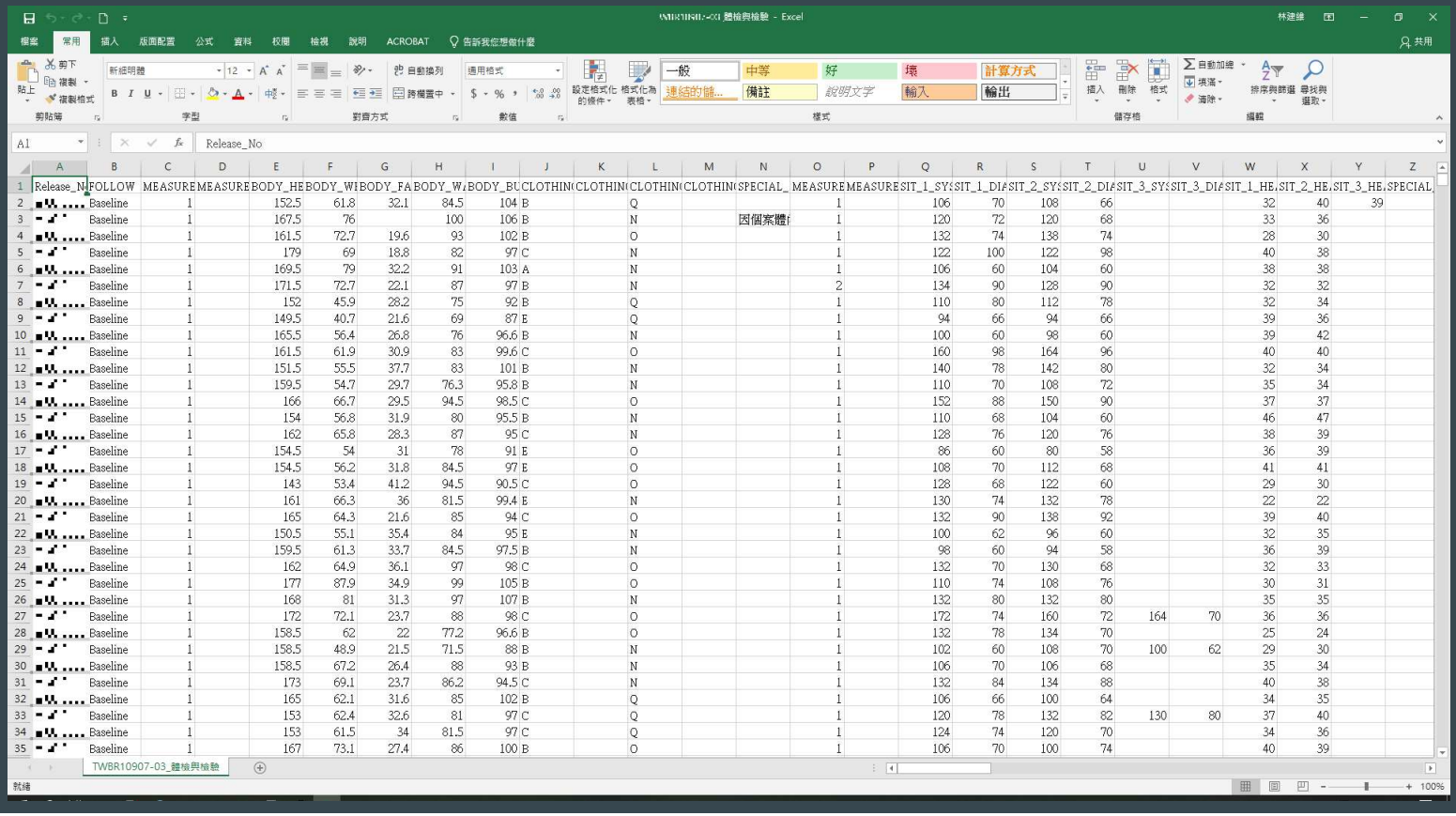
~3 million tubes

data

1. 一般參與者問卷資料：基本人口學變項、個人健康行為、生活環境、飲食狀況、家族疾病史、女性相關問題、經濟狀況、中醫體質
問項、簡易智能量表(下載詳細問卷內容PDF檔)(下載簡易版問卷內容PDF檔) (下載追蹤版問卷內容PDF檔) 。
2. 一般參與者身體檢測資料：身高、體重、體脂肪、腰臀圍、血壓、脈搏、骨密度、肺功能(下載詳細檢測內容PDF檔) 。

3. 一般參與者血液與尿液檢驗資料(下載詳細檢測內容PDF檔)(下載一般參與者血液與尿液檢驗資料項目PDF檔)：
(血液學檢驗項目)：紅血球、白血球、血小板、血紅素、血球比容、醣化血色素值。
(血清學檢驗項目)：飯前血糖、總膽固醇、三酸甘油脂、高密度脂蛋白膽固醇、低密度脂蛋白膽固醇。
(肝膽功能類檢驗項目)：總膽紅素、白蛋白、血清麩胺酸苯醋酸轉氨基酶、血清麩胺酸丙酮酸轉氨基酶、 γ-麩胺醯轉移酶、甲型胎兒血清蛋白。
(腎臟功能類檢驗、尿液檢驗)：血中尿素氮、肌酸酐、尿酸、尿中微白蛋白。
(病毒檢驗項目)：C型肝炎抗體、 B型肝炎表面抗原、 B型肝炎表面抗體、 B型肝炎核心抗體、 B型肝炎e抗原。
4. 一般參與者生物檢體：DNA、血漿、尿液。

dataland



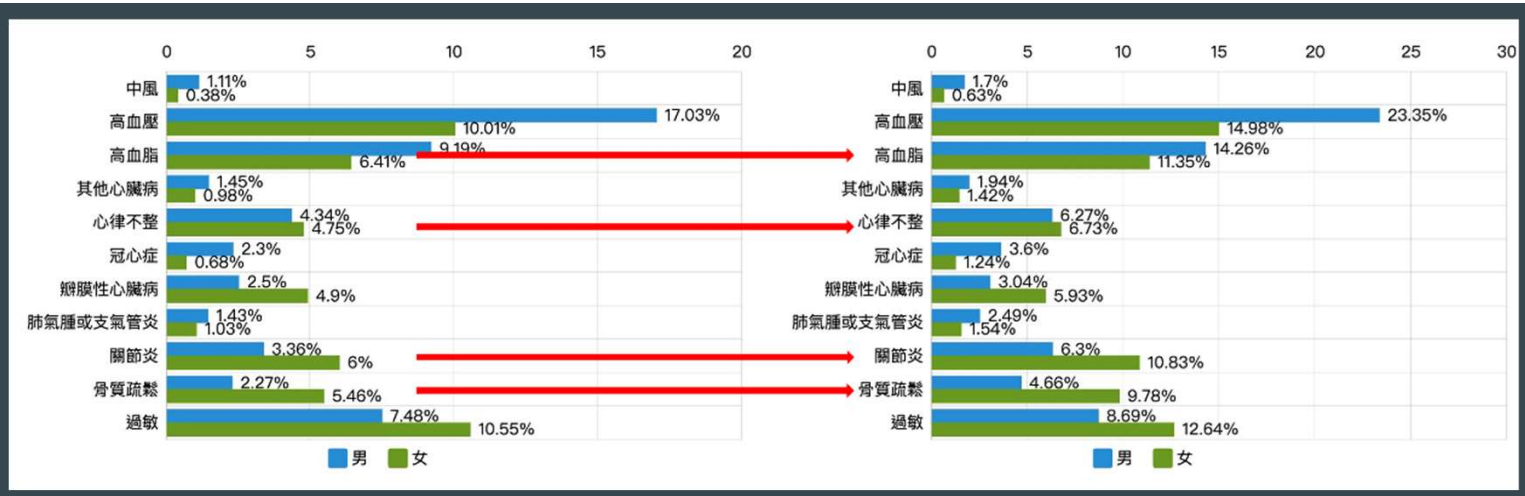
Release_N	FOLLOW	MEASURE	MEASURE	BODY_HE	BODY_WI	BODY_FA	BODY_WA	BODY_BU	CLOTHIN	CLOTHIN	CLOTHIN	CLOTHIN	SPECIAL_	MEASURE	MEASURE	SIT_1_SY	SIT_1_DI	SIT_2_SY	SIT_2_DI	SIT_3_SY	SIT_3_DI	SIT_1_HE	SIT_2_HE	SIT_3_HE	SPECIAL
Release_No																									
	Baseline	1		152.5	61.8	32.1	84.5	104	B		Q			1		106	70	108	66			32	40	39	
	Baseline	1		167.5	76		100	106	B		N		因個案體	1		120	72	120	68			33	36		
	Baseline	1		161.5	72.7	19.6	93	102	B		O			1		132	74	138	74			28	30		
	Baseline	1		179	69	18.8	82	97	C		N			1		122	100	122	98			40	38		
	Baseline	1		169.5	79	32.2	91	103	A		N			1		106	60	104	60			38	38		
	Baseline	1		171.5	72.7	22.1	87	97	B		N			2		134	90	128	90			32	32		
	Baseline	1		152	45.9	28.2	75	92	B		Q			1		110	80	112	78			32	34		
	Baseline	1		149.5	40.7	21.6	69	87	E		Q			1		94	66	94	66			39	36		
	Baseline	1		165.5	56.4	26.8	76	96.6	B		N			1		100	60	98	60			39	42		
	Baseline	1		161.5	61.9	30.9	83	99.6	C		O			1		160	98	164	96			40	40		
	Baseline	1		151.5	55.5	37.7	83	101	B		N			1		140	78	142	80			32	34		
	Baseline	1		159.5	54.7	29.7	76.3	95.8	B		N			1		110	70	108	72			35	34		
	Baseline	1		166	66.7	29.5	94.5	98.5	C		O			1		152	88	150	90			37	37		
	Baseline	1		154	56.8	31.9	80	95.5	B		N			1		110	68	104	60			46	47		
	Baseline	1		162	65.8	28.3	87	95	C		N			1		128	76	120	76			38	39		
	Baseline	1		154.5	54	31	78	91	E		O			1		86	60	80	58			36	39		
	Baseline	1		154.5	56.2	31.8	84.5	97	E		O			1		108	70	112	68			41	41		
	Baseline	1		143	53.4	41.2	94.5	90.5	C		O			1		128	68	122	60			29	30		
	Baseline	1		161	66.3	36	81.5	99.4	E		N			1		130	74	132	78			22	22		
	Baseline	1		165	64.3	21.6	85	94	C		O			1		132	90	138	92			39	40		
	Baseline	1		150.5	55.1	35.4	84	95	E		N			1		100	62	96	60			32	35		
	Baseline	1		159.5	61.3	33.7	84.5	97.5	B		N			1		98	60	94	58			36	39		
	Baseline	1		162	64.9	36.1	97	98	C		O			1		132	70	130	68			32	33		
	Baseline	1		177	87.9	34.9	99	105	B		O			1		110	74	108	76			30	31		
	Baseline	1		168	81	31.3	97	107	B		N			1		132	80	132	80			35	35		
	Baseline	1		172	72.1	23.7	88	98	C		O			1		172	74	160	72	164	70	36	36		
	Baseline	1		158.5	62	22	77.2	96.6	B		O			1		132	78	134	70			25	24		
	Baseline	1		158.5	48.9	21.5	71.5	88	B		N			1		102	60	108	70	100	62	29	30		
	Baseline	1		158.5	67.2	26.4	88	93	B		N			1		106	70	106	68			35	34		
	Baseline	1		173	69.1	23.7	86.2	94.5	C		N			1		132	84	134	88			40	38		
	Baseline	1		165	62.1	31.6	85	102	B		Q			1		106	66	100	64			34	35		
	Baseline	1		153	62.4	32.6	81	97	C		Q			1		120	78	132	82	130	80	37	40		
	Baseline	1		153	61.5	34	81.5	97	C		Q			1		124	74	120	70			34	36		
	Baseline	1		167	73.1	27.4	86	100	B		O			1		106	70	100	74			40	39		

data

			單位換算：GB	申請人數	所需儲存容量
Genotyping					
.CEL檔		30 MB	0.03	114601	3438.03
NGS sequencing					
VCF	GATK	5 GB	5	1496	7480
	Ion Proton	500 MB	0.5	514	257
MET sequencing					
.txt與.idat圖檔		100 MB	0.1	2474	247.4
HLA typing					
.csv與BAM檔		400 MB	0.4	1102	440.8
Metabolomics					
FID與SER檔		15 MB	0.015	869	13.035
總計（GB）					**11876.265**
總計（TB）					**11.876265**

			單位換算：GB	申請人數	所需儲存容量
NGS sequencing					
Fastq（.gz，Only Illumina）		80 GB	80	1496	119680
VCF	iSAAC	5 GB	5	996	4980
BAM	iSAAC	50 GB	50	996	49800
	GATK（BWA）	150 GB	150	1496	224400
	Ion Proton	400 GB	400	514	205600
總計（GB）					**604460**
總計（TB）					**604.46**

	單位換算：GB	申請人數	所需儲存容量
Genotyping			
PLINK Imputation (TWB_1.0、2.0、1.0+2.0)		**數位資料集**	**2TB**
進階追蹤資料		**數位資料集**	**3TB**
joint calling		Illumina	**500G**

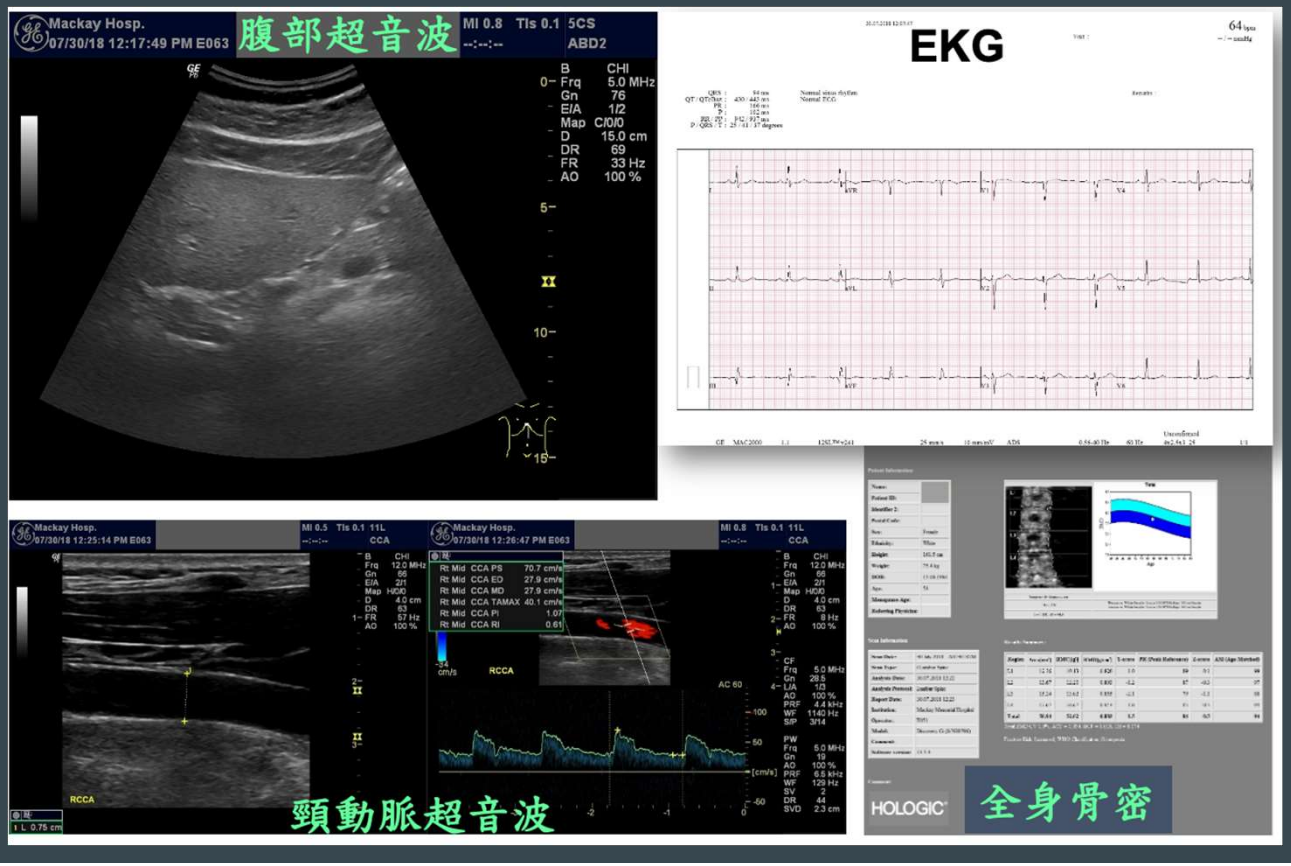


Urinary metabolites (μg/g crea.)	TWB (n=1,155)	NHANES (n=2,974)[a]
Melamine (μg/mmol crea.)	0.46 (0.43-0.49)	---
MEHP	11.37 (10.69-12.09)	N.D.[b]
MEOHP	8.17 (7.81-8.54)	3.70 (3.49-3.92)
MEHHP	12.68 (12.06-13.33)	5.86 (5.62-6.12)
MECPP	19.17 (18.36-20.02)	9.14 (8.58-9.74)
MCMHP	3.88 (3.61-4.16)	--
MBzP	1.04 (0.98-1.1)	4.63 (4.06-5.28)
MnBP	20.51 (19.46-21.61)	10.2 (9.53-10.9)
MiBP	8.38 (7.92-8.87)	8.71 (8.06-9.42)
MEP	12.86 (11.82-13.99)	34.7 (31.0-39.0)
MMP	2.22 (2.09-2.36)	---
MiNP	N.D.[b]	N.D.[b]

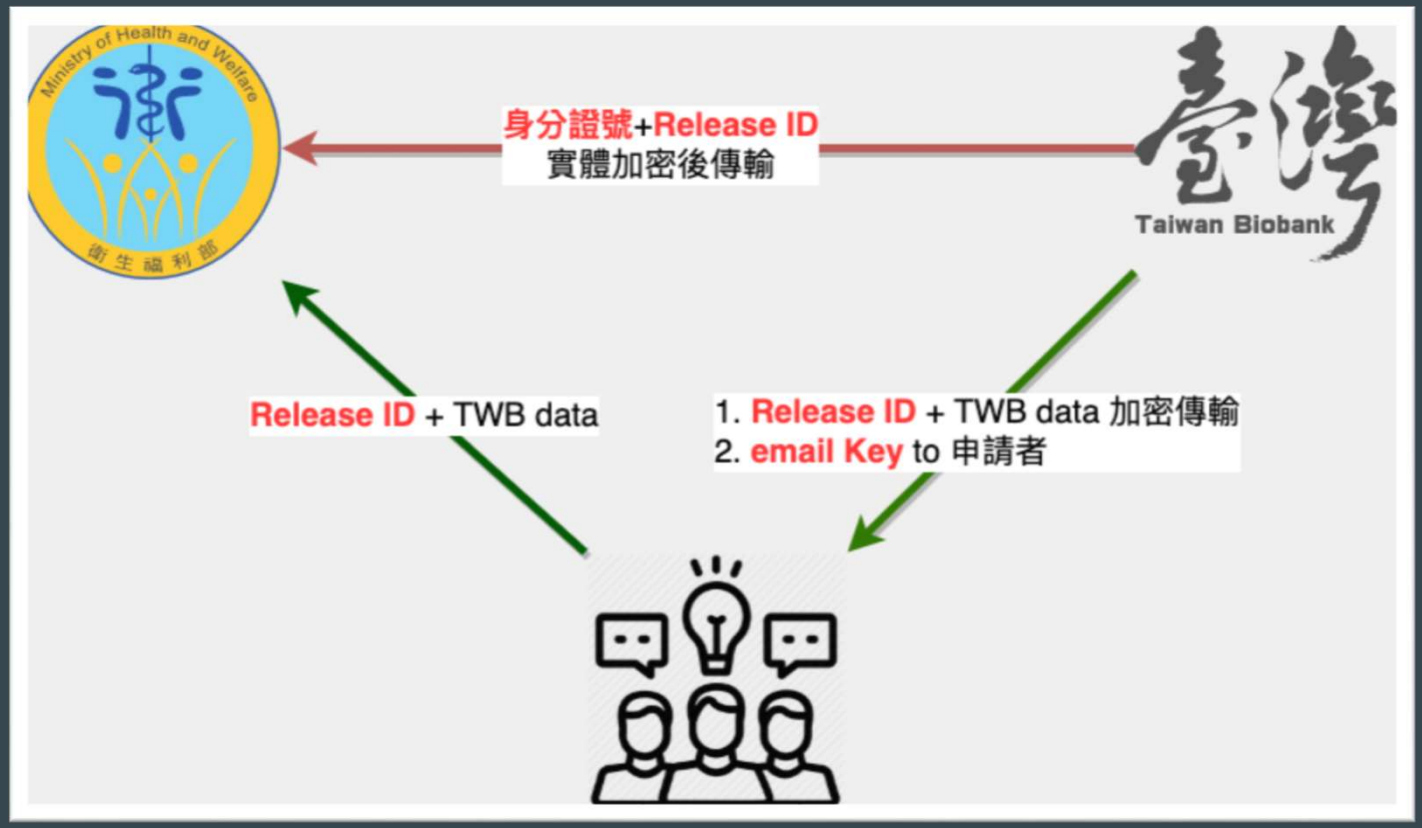
[a]Data from the urine samples of 2015-2016 for the U.S. population from the National Health and Nutrition Examination Survey. (NHANES, 2019)
[b]N.D., the chemical was analyzed but the proportion of results below limit of detection was too high to provide a valid result.

Characteristics	Male (%)	female (%)
BMI		
<18.5	1.29	4.42
18.5-24	35.83	56.96
24-27	35.25	22.58
>27 (Obesity)	27.62	16.05
Waist-Hip Ratio		
M ≥ 0.92; F ≥ 0.88	36.05	30.00
Body Fat Rate (male, female)		
≦17% , ≦20%	12.42	2.02
17-23% , 20-27%	39.11	20.02
23-25%, 27-30%	16.02	17.81
>25% , >30%	32.45	60.11

advance follow-up program

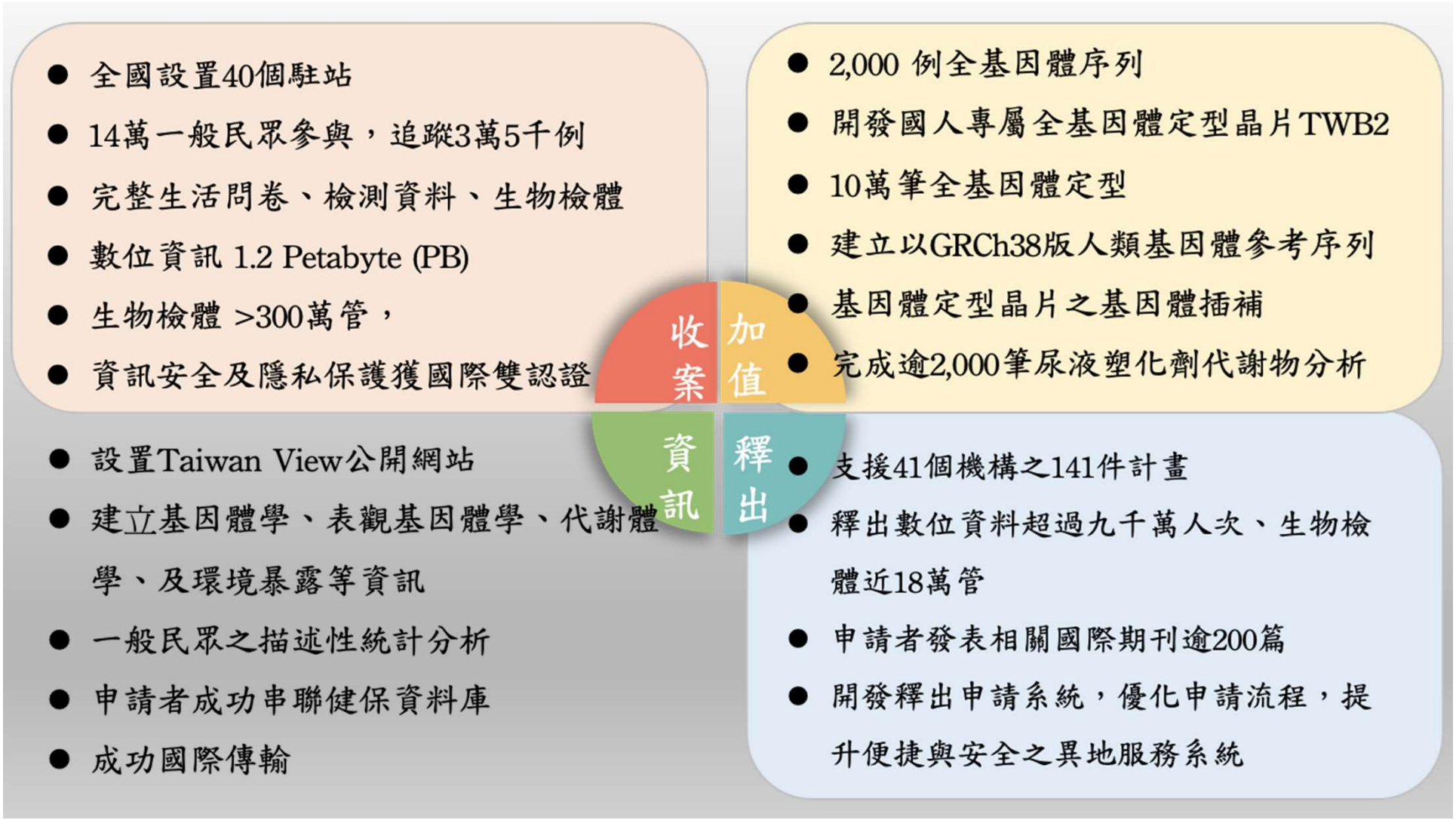


TWB-NHIRD


Ministry of Health and Welfare
衛生福利部
身分證號+Release ID
實體加密後傳輸
臺灣
Taiwan Biobank
Release ID + TWB data
1. Release ID + TWB data 加密傳輸
2. email Key to 申請者

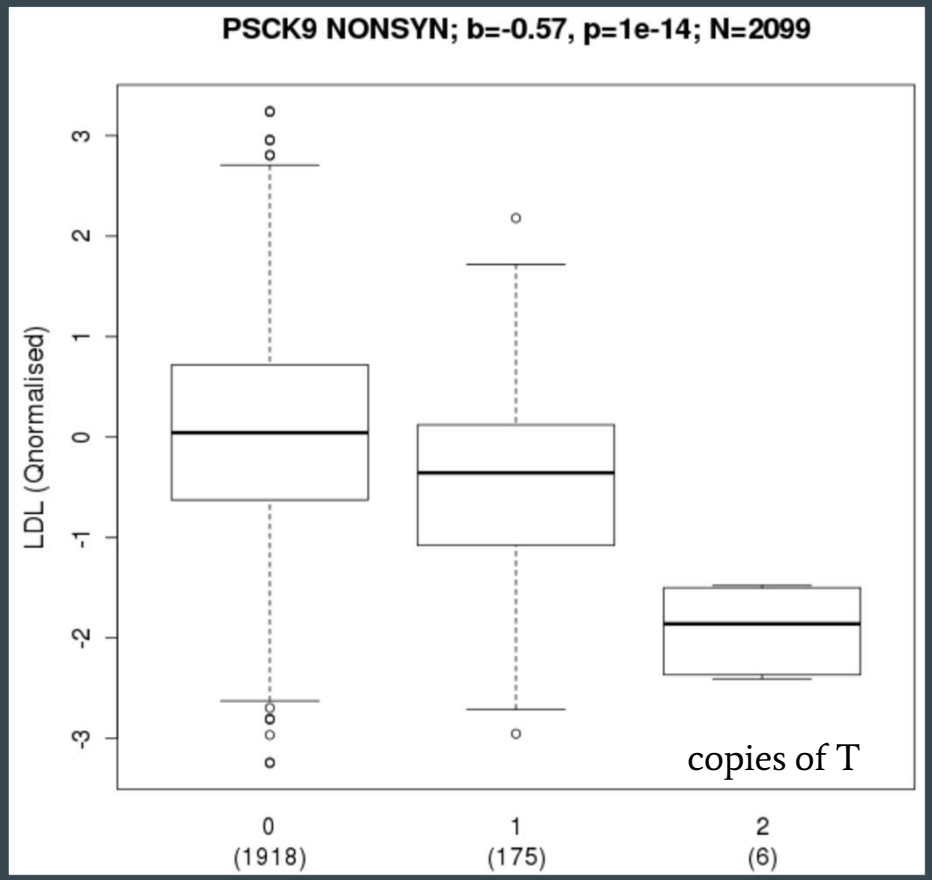
TWB - NHIRD

Diseases	Sex	TWB (%)	NHIRD (%)
Diabetes	Male	10.98	12.53
	Female	7.25	8.84
Hypertension	Male	29.12	26.7
	Female	16.86	17.08
Hyperlipidemia	Male	26.69	26.09
	Female	20.51	21.05


收案
加值
資訊
釋出
全國設置40個駐站
14萬一般民眾參與，追蹤3萬5千例
完整生活問卷、檢測資料、生物檢體
數位資訊 1.2 Petabyte (PB)
生物檢體 >300萬管，
資訊安全及隱私保護獲國際雙認證
2,000 例全基因體序列
開發國人專屬全基因體定型晶片TWB2
10萬筆全基因體定型
建立以GRCh38版人類基因體參考序列
基因體定型晶片之基因體插補
完成逾2,000筆尿液塑化劑代謝物分析
設置Taiwan View公開網站
建立基因體學、表觀基因體學、代謝體學、及環境暴露等資訊
一般民眾之描述性統計分析
申請者成功串聯健保資料庫
成功國際傳輸
支援41個機構之141件計畫
釋出數位資料超過九千萬人次、生物檢體近18萬管
申請者發表相關國際期刊逾200篇
開發釋出申請系統，優化申請流程，提升便捷與安全之異地服務系統

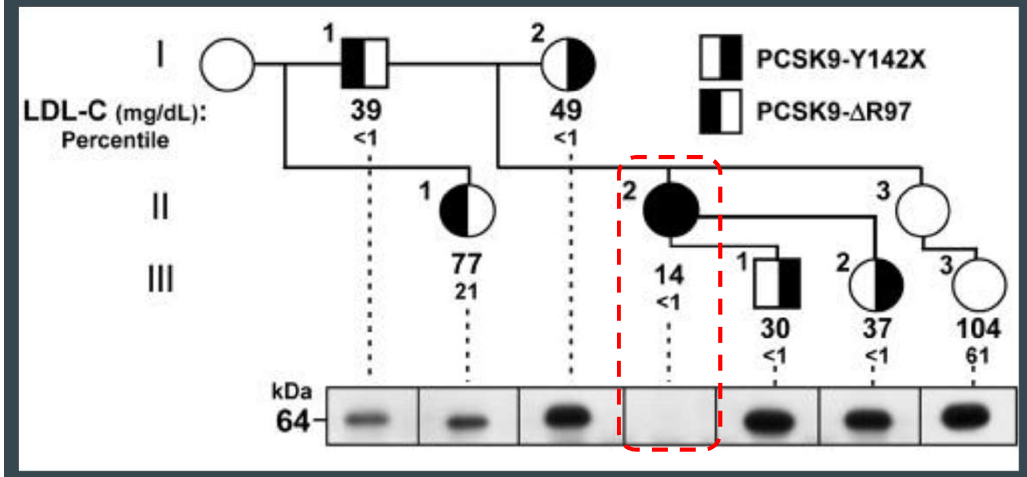

Success story

GENETIC VARIANT “RS11591147” IN PCSK9



- Carriers of T variant have lower levels of LDL cholesterol than carriers of G variant
- LDL is a strong risk factor for heart disease

A HUMAN KNOCK-OUT OF PCSK9 (2006)



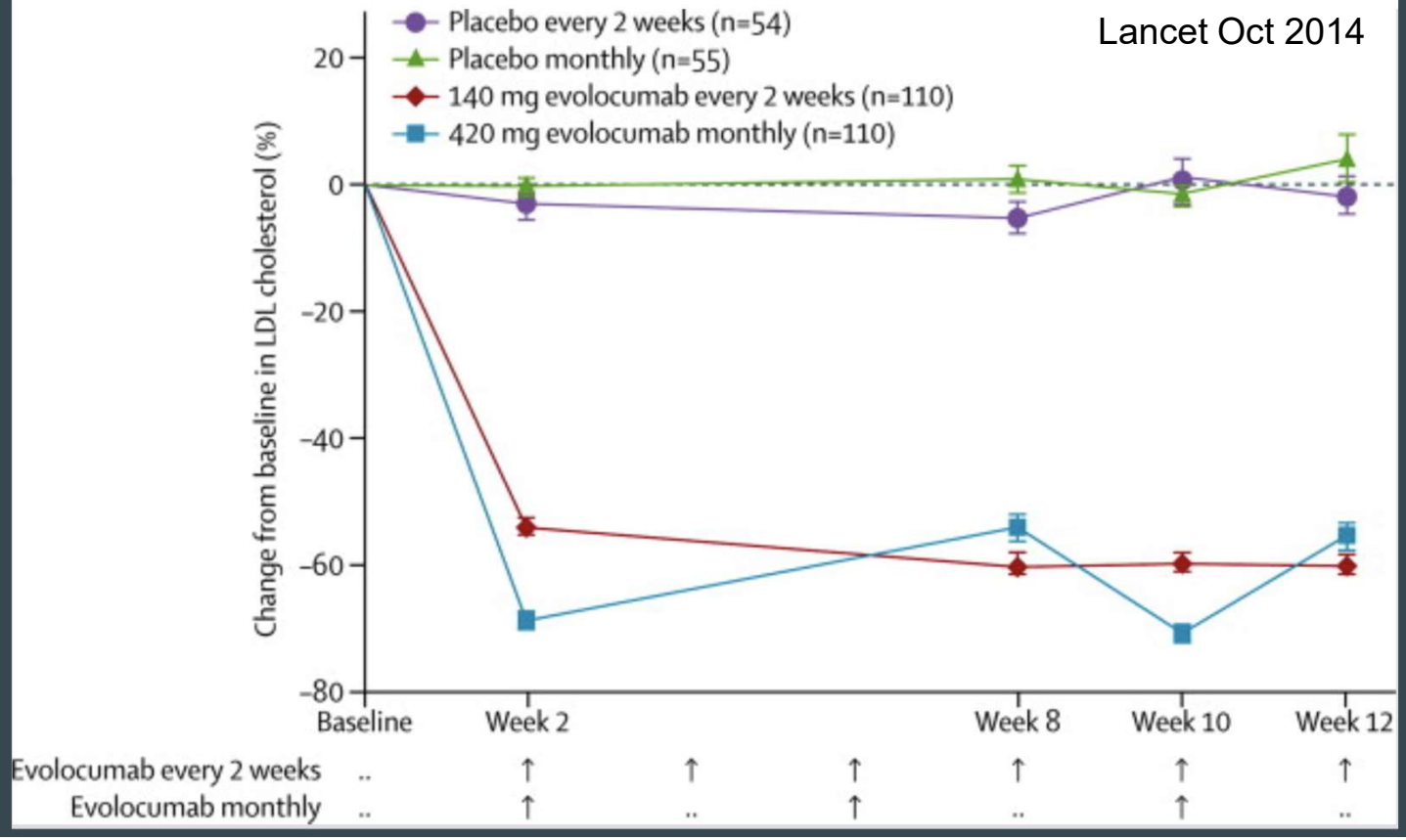
Individual II.2 has zero working copies of PCSK9 gene

- no circulating PCSK9 and an LDL-C of only 14 mg/dL
- apparently healthy, fertile, normotensive, college-educated woman with normal liver and renal function tests who works as an aerobics instructor
- Why is this very interesting observation? Inhibiting PCSK9 might be a safe way to reduce LDL

Zhao et al. AJHG 2006

PCSK9 inhibition with evolocumab (AMG 145) in heterozygous familial hypercholesterolaemia (RUTHERFORD-2): a randomised, double-blind, placebo-controlled trial

Lancet Oct 2014



ORIGINAL ARTICLE

N Engl J Med 2017; 376:1713-1722

Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease

Marc S. Sabatine, M.D., M.P.H., Robert P. Giugliano, M.D., Anthony C. Keech, M.D., Narimon Honarpour, M.D., Ph.D., Stephen D. Wiviott, M.D., Sabina A. Murphy, M.P.H., Julia F. Kuder, M.A., Huei Wang, Ph.D., Thomas Liu, Ph.D., Scott M. Wasserman, M.D., Peter S. Sever, Ph.D., F.R.C.P., and Terje R. Pedersen, M.D. for the FOURIER Steering Committee and Investigators*

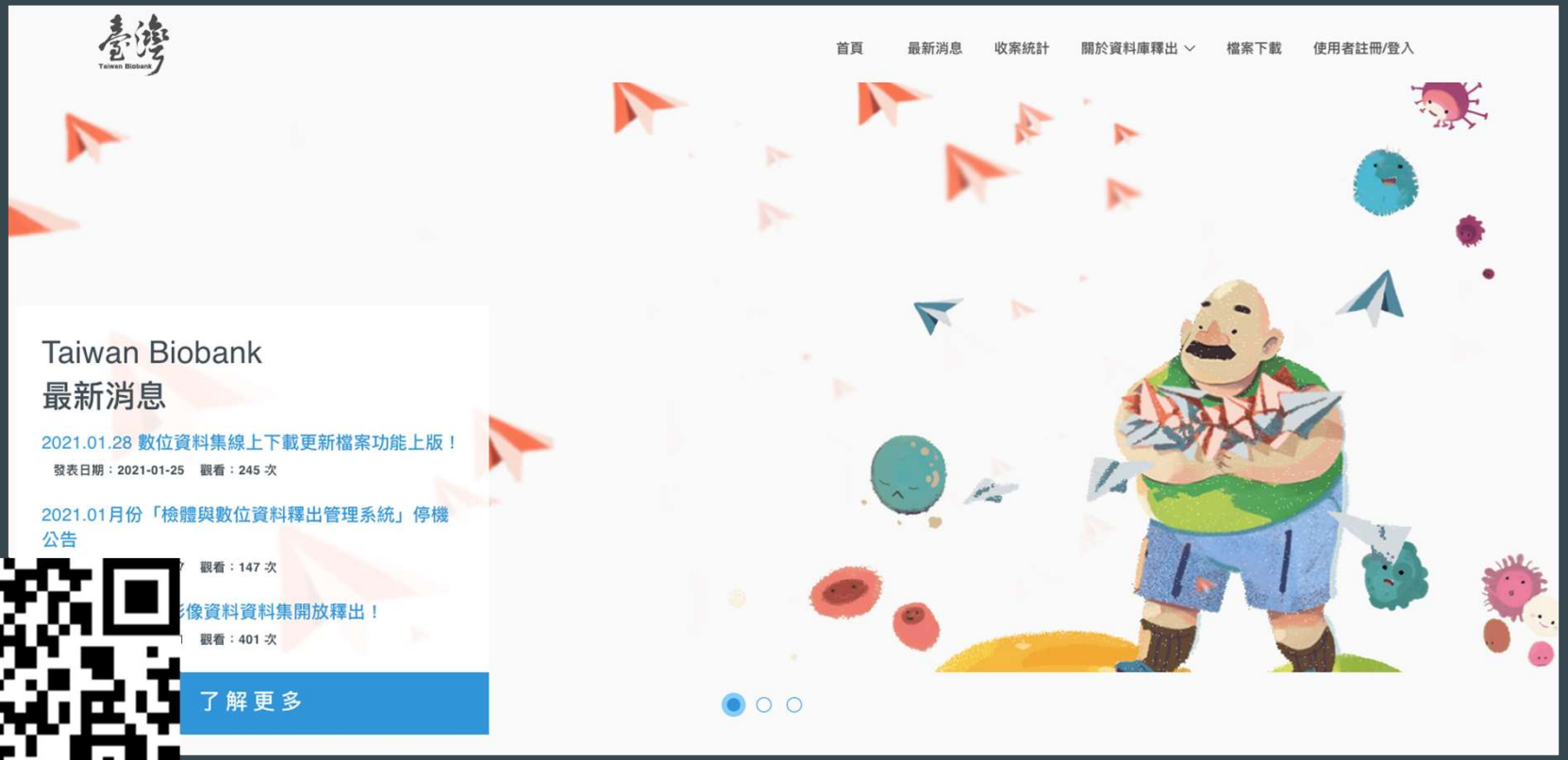
FDA Approves Amgen's Repatha (evolocumab) to Prevent Heart Attack and Stroke

Dec 1 2017

In the Repatha cardiovascular outcomes study (FOURIER), Repatha reduced the risk of heart attack by 27%, the risk of stroke by 21% and the risk of coronary revascularization by 22%..


Let's write !!

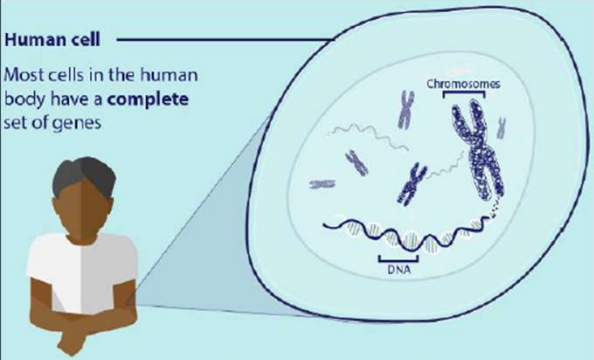
https://www.biobank.org.tw/



https://reurl.cc/OXkXM7

Human cell

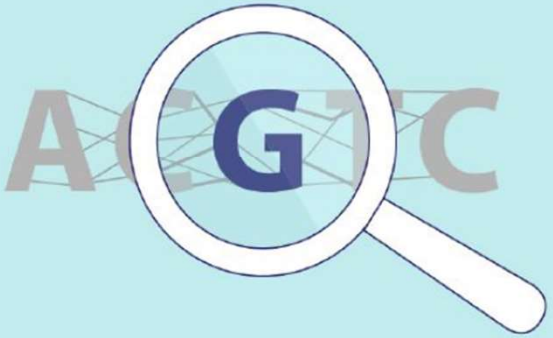
Most cells in the human body have a **complete** set of genes



Your **genome** is one whole set of all your genes plus all the DNA between your genes.

There are around **20,000** genes in your **genome**

Now



We know that the non-gene (non-coding) parts of your genome may have a role to play so we look at the whole thing, every single letter, and how the different parts work together.

5%

were analysed

Only the active genes were looked at

Your genome

with

3 Billion

pairs of letters in the human genome

AAGTAATATGC
TTCTAGGCGTC
TCAAGATGCAT
CTAGCACAGC
GCCCTTTATTA
TCTCTATACTCA
ACTACTAGGGC
TATTTCATATCT
AAATACGCTCG
AGGCTACTGAC
TTATGCTATCG
ATCTCGAGCGC
TDCCGTAATTT
TCGCGAATCAG
AAGTAATATGC
TTCTAGGCGTC
TCAAGATGCAT
CTAGCACAGC
GCCCTTTATTA

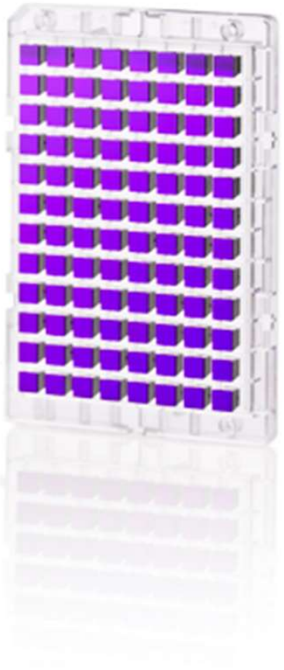
95%

were unused

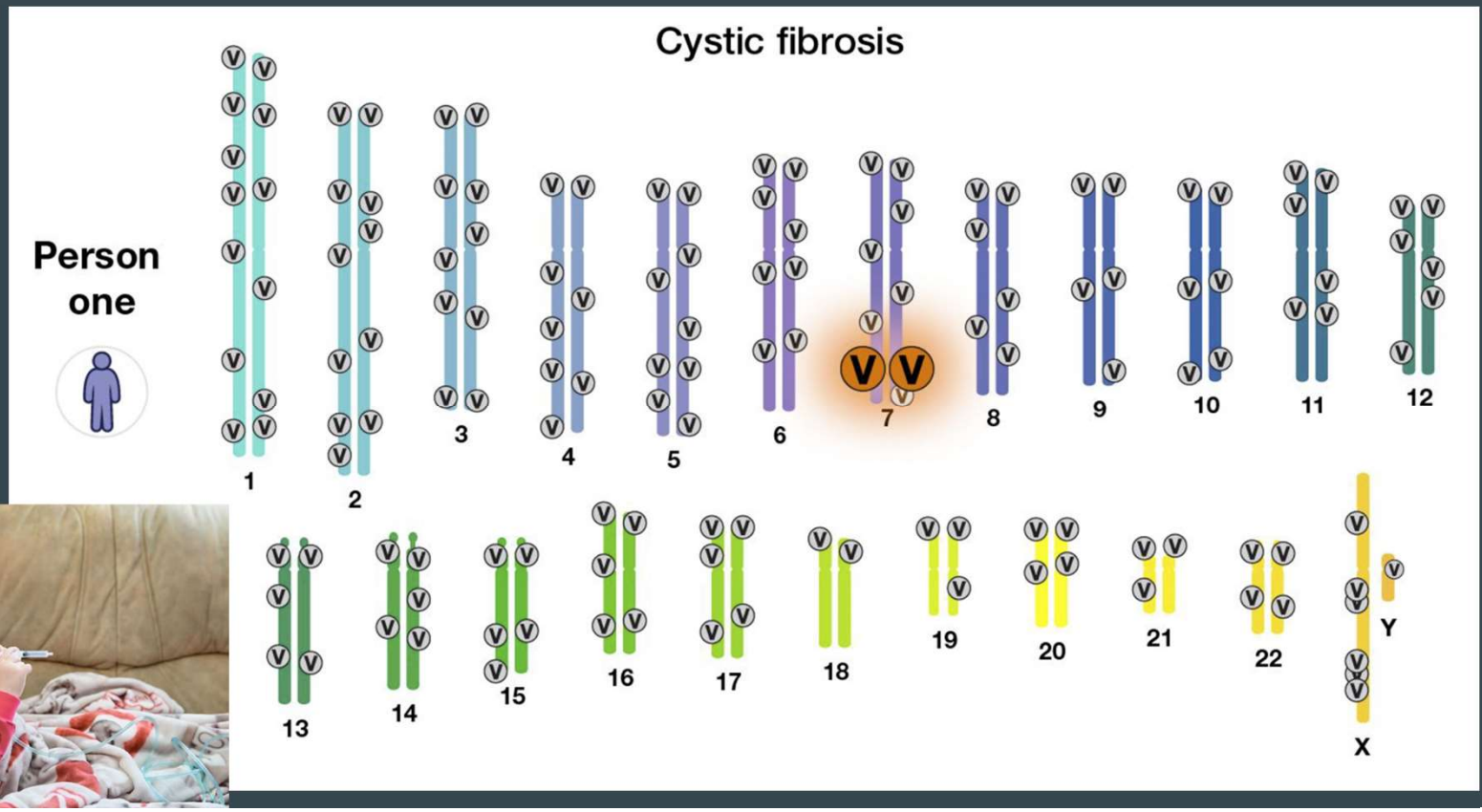
All the non-gene sections that we didn't understand were disregarded as useless

BIOBANK • 健康世代

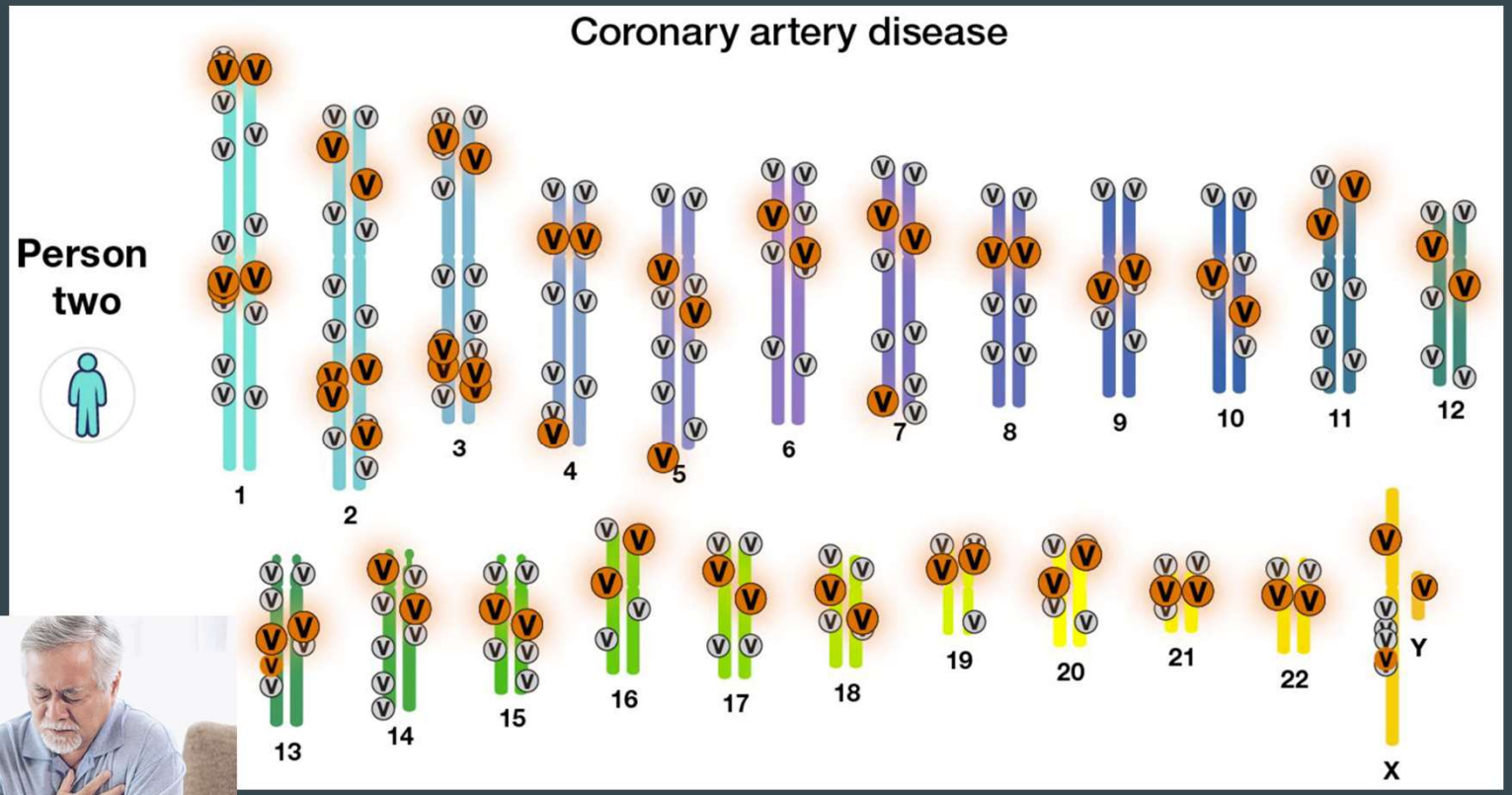
TWB2 axiom array

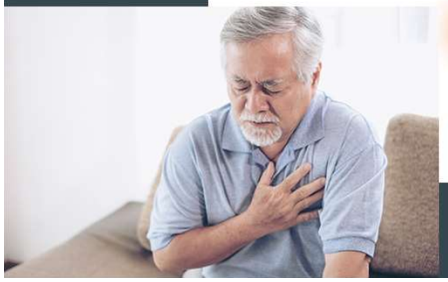


cystic fibrosis (囊狀纖維化) → transmembrane conductance regulator (CFTR) gene on chromosome 7.



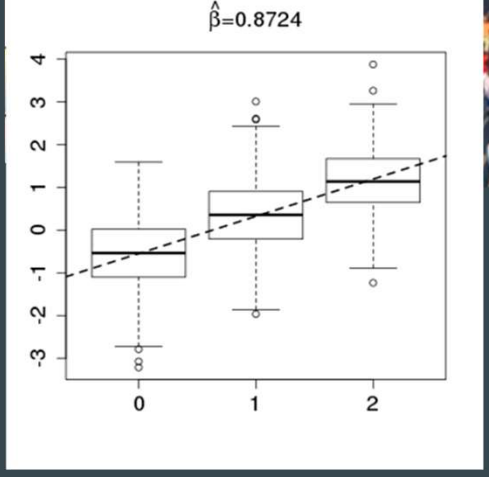
complex diseases (polygenic disease)

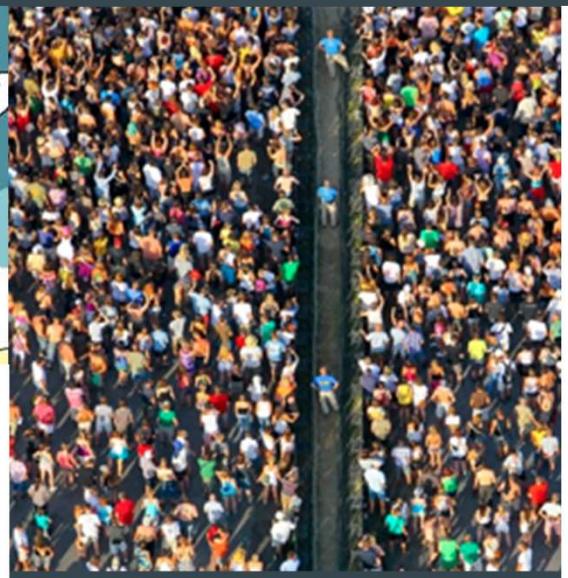




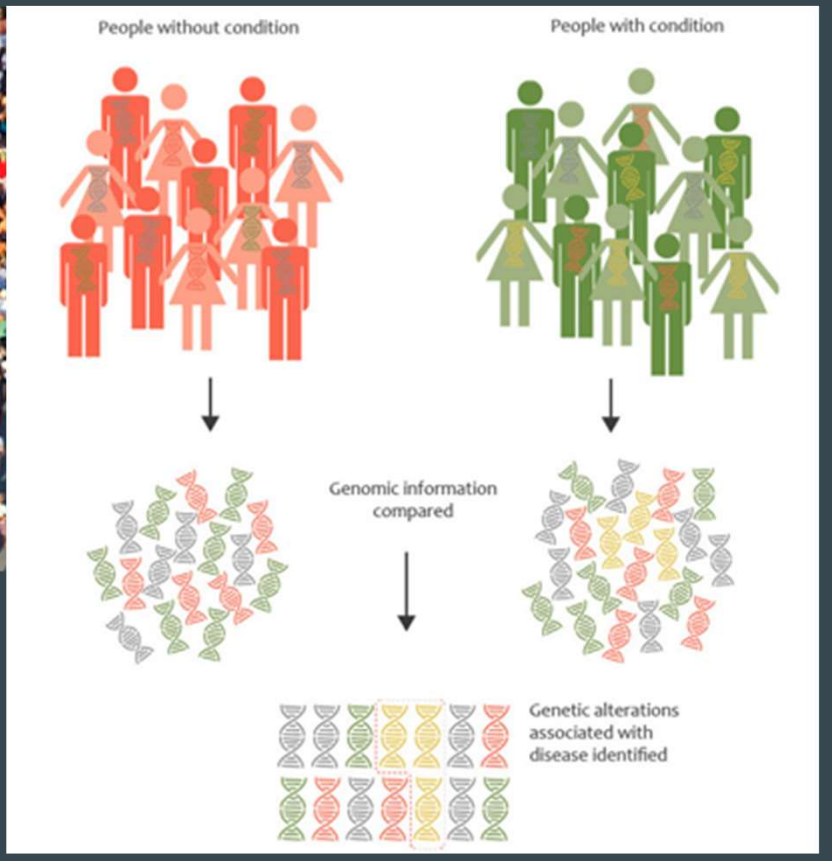
GWAS


GOD... THE HUMAN GENOME CODE'S BEEN UNRAVELLED
DAMN HACKERS!!! NOW, I HAVE TO CHANGE THE PASSWORD

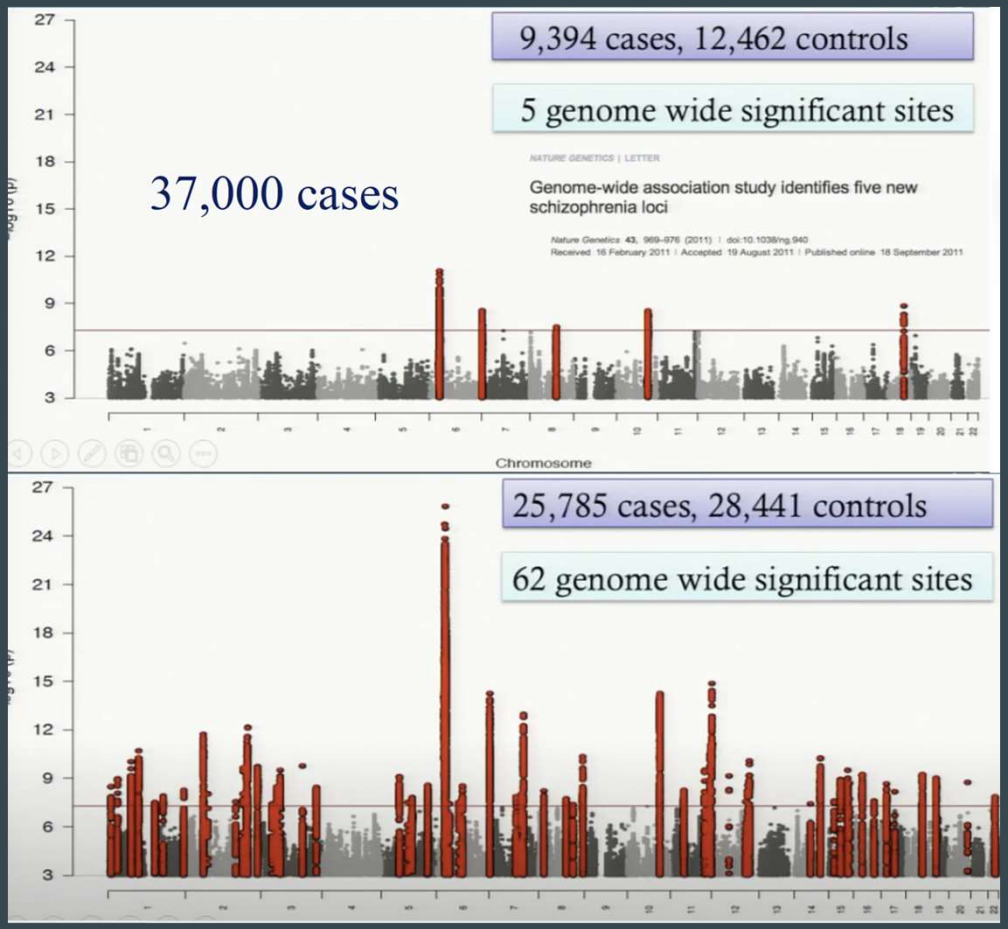

$\hat{\beta}$=0.8724
4
3
2
1
0
-1
-2
-3
0
1
2

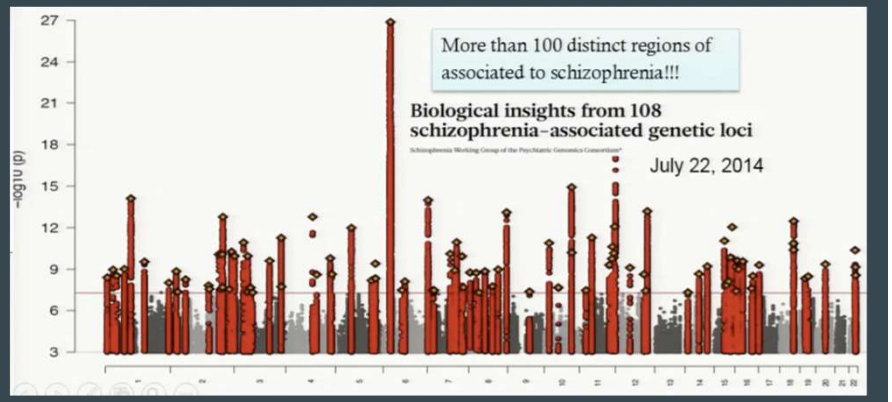



6. ADVAIR DISKUS (fluticasone propionate)
Asthma


People without condition
People with condition
Genomic information compared
Genetic alterations associated with disease identified

the power of sample size - schizophrenia | psychiatric genomics consortium





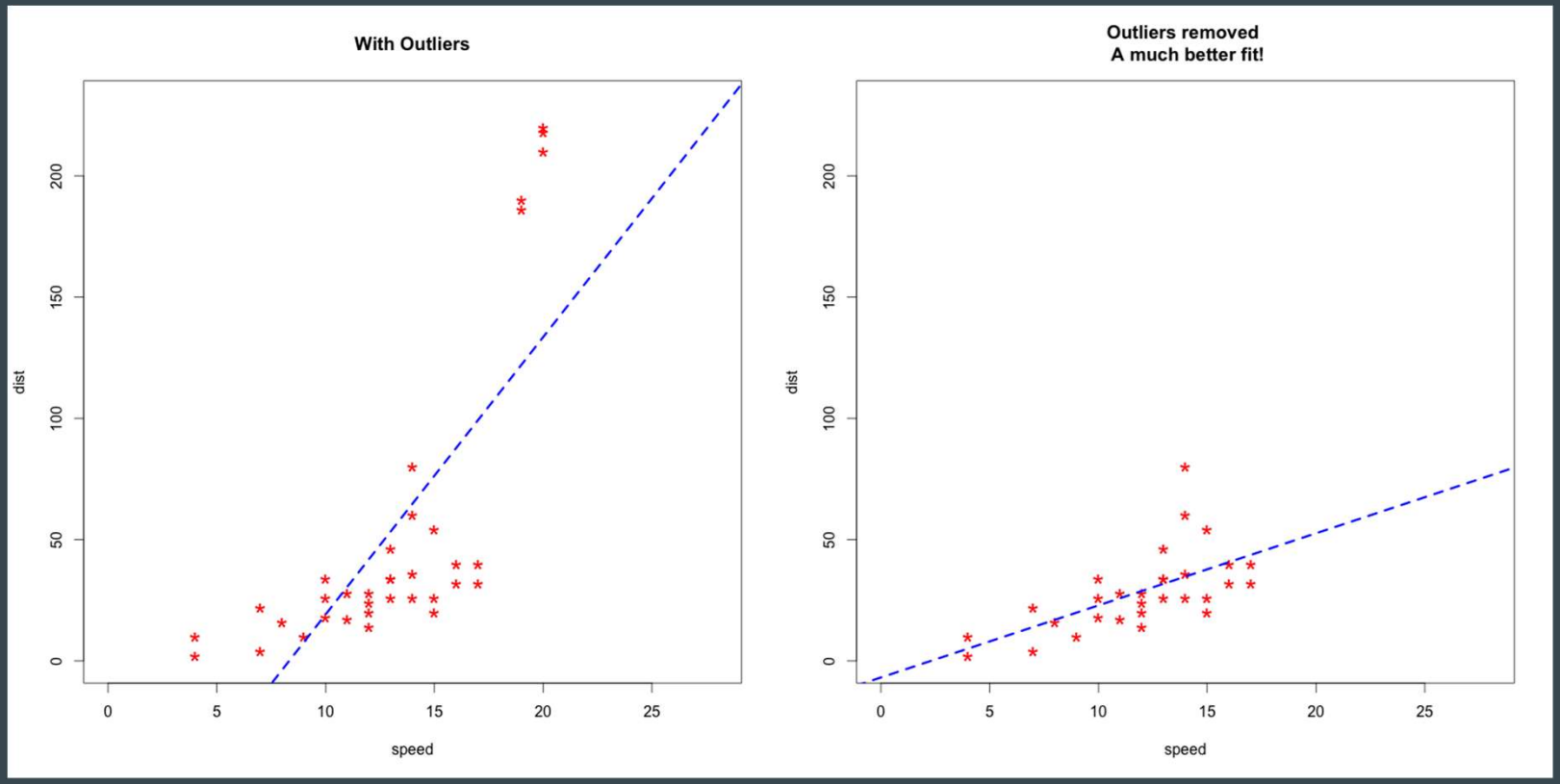
	用藥	未用藥
男	**20**	**40**
女	**40**	**20**

用藥組生病人數：男 8 人，女 3 人
未用藥組生病人數：男 12 人，女 1 人

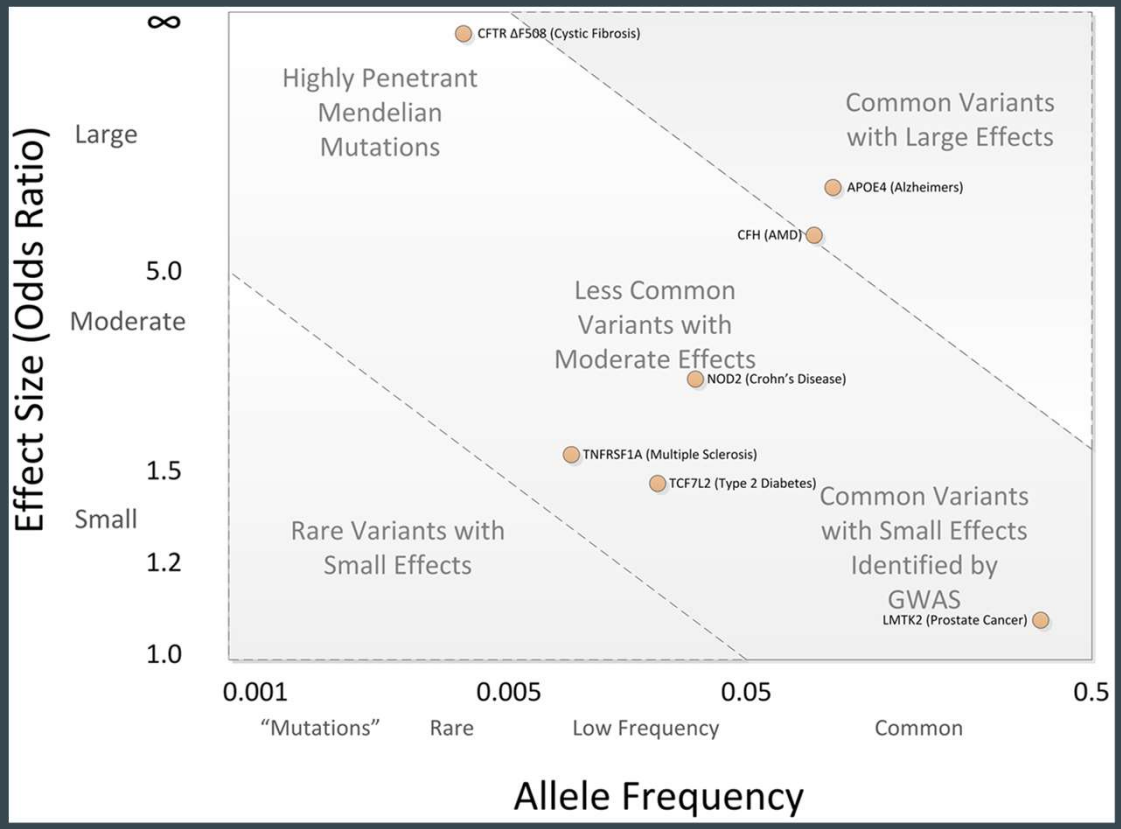
用藥組罹病率：　男 8/20 = 0.4， 女 3/40 = 0.075
未用藥組罹病率：男 12/40 = 0.3，女 1/20 = 0.05
結論：藥物對男性有害也對女性有害

用藥組罹病率：(8+3)/(20+40) = 0.18
未用藥組罹病率：(12+1)/(40+20) = 0.22
結論：藥物對人有益

linear regression (linearMod <- lm(glucose ~ bmi, data=twb))



EFFECT, MAF, AND REGION OF POWER



https://doi.org/10.1371/journal.pcbi.1002822


THE ROAD TO
SUCCESS
IS ALWAYS
UNDER
CONSTRUCTION

Published Ahead of Print on December 2, 2019 as 10.1212/WNL.0000000000008700

ARTICLE

NOTCH3 cysteine-altering variant is an important risk factor for stroke in the Taiwanese population

Yi-Chung Lee, MD, PhD, Chih-Ping Chung, MD, PhD, Ming-Hong Chang, MD, Shuu-Jiun Wang, MD, PhD, and Yi-Chu Liao, MD, PhD

Correspondence
Dr. Liao
ycliao5@vghtpe.gov.tw

Neurology® 2020;94:1-10. doi:10.1212/WNL.0000000000008700

NOTCH3 cysteine-altering variant is an important risk factor for stroke in the Taiwanese population

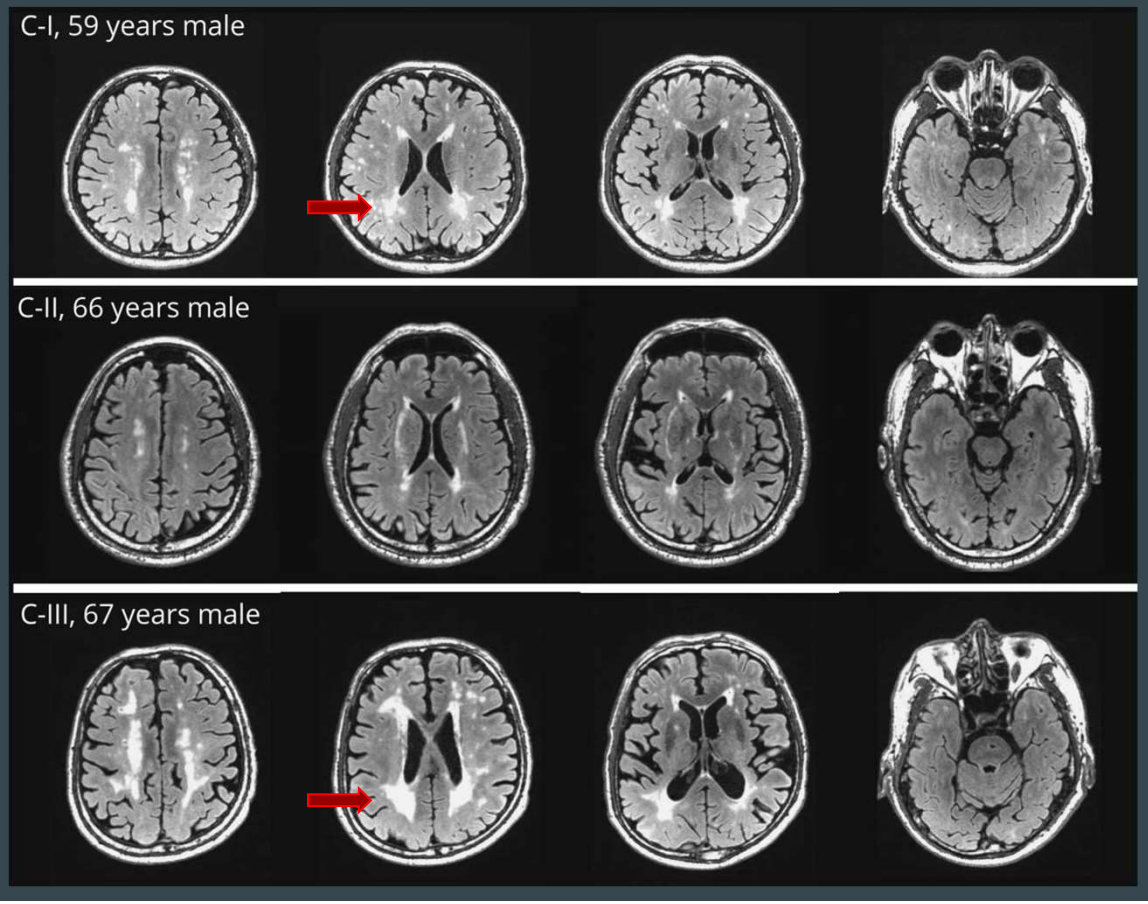
We queried the Taiwan Biobank database for cysteine-altering mutations in exons 2–24 of NOTCH3 within these genomes. The reference coding sequence of NOTCH3, NM_000435.3, was used for annotating the variants. (p.R544C (c.1630G>A), p.C853Y (c.2558G>A), and p.C884Y (c.2651G>A))

The cysteine-altering NOTCH3 variants identified from the Taiwan Biobank database were genotyped in the control participants and patients with stroke using the TaqMan genotyping assay

Only the NOTCH3 p.R544C variant was found in 4 individuals (TP-VGH (n =550))

	Controls			Patients with stroke		
	Taiwan biobank (n = 6,488)	TP-VGH (n = 550)	Biobank + TP-VGH (n = 7,038)	TP-VGH (n = 350)	TC-VGH (n = 450)	TP-VGH + TC-VGH (n = 800)
Male	2,293 (35.3)	230 (41.8)	2,523 (35.8)	248 (70.9)	327 (72.7)	575 (71.9)
Age, y	48.4 ± 10.9	56.7 ± 15.0	49.1 ± 11.5	64.4 ± 13.6	67.5 ± 13.2	66.2 ± 13.5
Hypertension	682 (10.6)	186 (33.8)	868 (12.4)	248 (70.9)	329 (73.1)	577 (72.1)
Diabetes	300 (4.6)	72 (13.1)	372 (5.3)	130 (37.1)	168 (37.3)	298 (37.3)
Hyperlipidemia	388 (6.0)	138 (25.1)	526 (7.5)	129 (36.9)	181 (40.2)	310 (38.8)
Smoking habit	1,228 (18.9)	98 (17.8)	1,326 (18.8)	122 (34.9)	136 (30.4)	258 (32.4)
Alcohol consumption	387 (6.0)	108 (19.6)	495 (7.0)	56 (16.0)	87 (23.2)	143 (19.7)
Family history of stroke	1,146 (17.8)	—	—	40 (14.1)	14 (4.0)	54 (8.5)
***NOTCH3* p.R544C mutation (+)**	56 (0.9)	4 (0.7)	60 (0.9)	8 (2.3)	9 (2.0)	17 (2.1)

Physical examination revealed that they were free of neurologic deficits. Three of them received brain MRI scans, and all had a variable degree of leukoencephalopathy

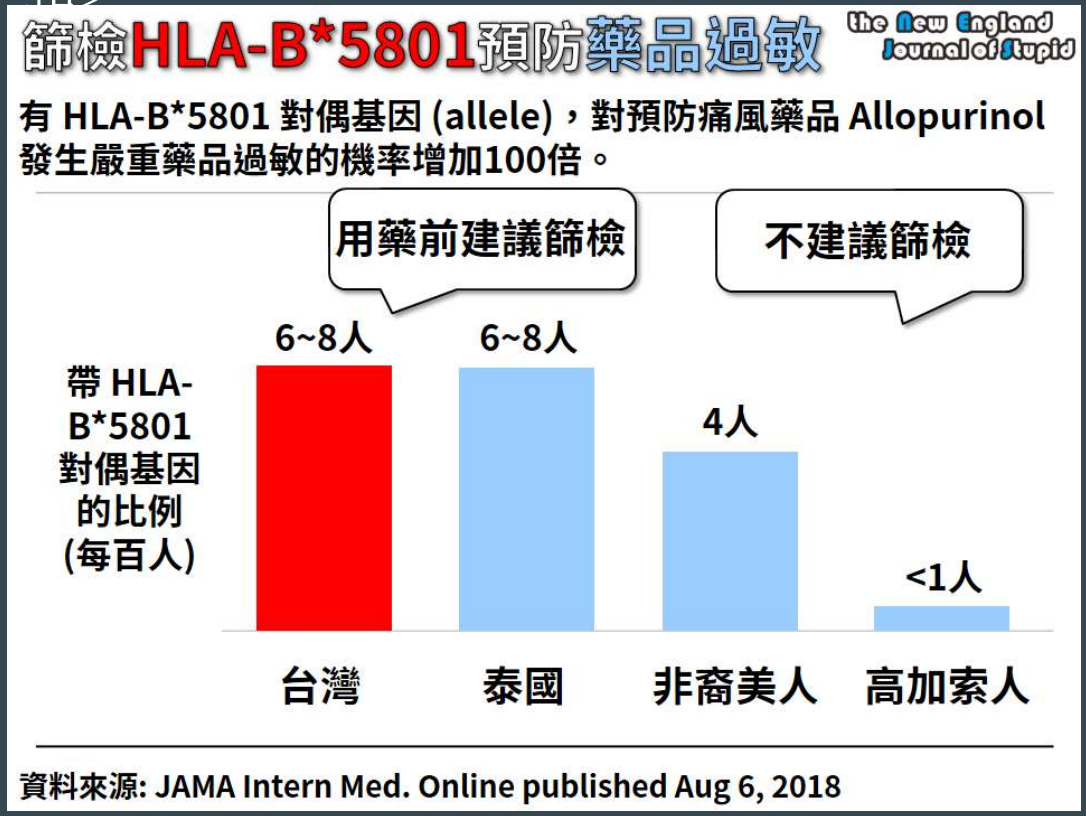


To clarify the role of these cysteine-altering NOTCH3 mutations in ischemic stroke

genotyped 800 patients with ischemic stroke(缺血性腦中風). No single patient was found to carry NOTCH3 p.C853Y or p.C884Y.

after that...

國內醫院 基因體檢測門診



http://jerryljw.blogspot.com/2018/08/hla-b5801-use-of-hla-b5801-genotyping.html

frequency distribution of pharmacogenetic phenotypes predicted by genotypes of TWB cohort

Gene	Drug	Rx[a]/year	EM	IM	PM	ADR[b] carrier rate
CYP2B6	Efavirenz	1,662,525	66.0%	30.5%	3.6%	
CYP2C19	Clopidogrel	63,664,076	39.8%	56.4%	3.8%	
CYP2C9	Celecoxib	65,058,810	93.6%	6.3%	0.1%	
CYP3A5	Tacrolimus	10,272,406	8.1%	40.6%	51.2%	
IL28	Peginterferon	40,941	88.6%	11.1%	0.3%	
NAT2	Isoniazid	7,885,251	28.8%	59.2%	12.0%	
SLCO1B1	Simvastatin	50,695,934	78.9%	19.9%	1.3%	
TPMT	Azathioprine	7,435,217	97.0%	2.9%	0.02%	
UGT1A1	Atazanavir	719,793	53.2%	39.8%	7.0%	
VKORC1	Warfarin	16,121,944	1.1%	19.2%	79.7%	
HLA-A*3101	Carbamazepine	17,078,849				2.0%
HLA-B*1502	Carbamazepine	17,078,849				4.1%
HLA-B*5701	Abacavir	3,049,217				0.2%
HLA-B*5801	Allopurinol	23,888,472				10.5%
MT-RNR1	Amikacin	321,561				4.7%

[a]Rx = prescriptions.

[b]*ADR* = adverse drug reactions.

NPJ Genom Med. 2021 Feb 11;6(1):10.

polygenic risk score

Discovery GWAS

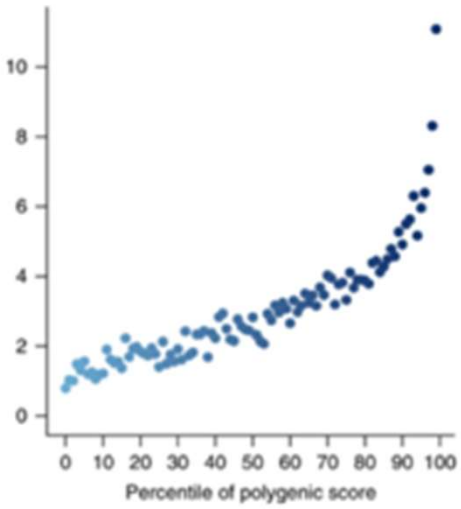
	Weight*	Risk Allele
SNP1	0.2	A
SNP2	-0.3	C
SNP3	0.1	G

Individual	Alleles SNP1	Alleles SNP2	Alleles SNP3
1	AT	AA	CG
2	AA	CA	GG
3	TT	AC	CG
4	TT	AA	GG
5	TA	CA	GC
6	AT	CA	CG
7	AA	AA	GG
8	AA	CC	CG
9	TA	CC	GC
10	AT	AA	CG

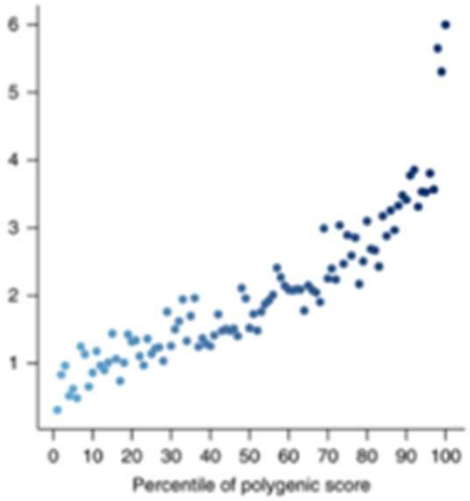
PRS:

Individual	SNP 1	SNP 2	SNP 3	**PRS**
1	0.2+0.0	0.0+0.0	0.0+0.1	**0.3**
2	0.2+0.2	-0.3+0.0	0.1+0.1	**0.3**
3	0.0+0.0	0.0-0.3	0.0+0.1	**-0.2**
4	0.0+0.0	0.0+0.0	0.1+0.1	**0.2**
5	0.0+0.2	-0.3+0.0	0.1+0.0	**0.0**
6	0.2+0.0	-0.3+0.0	0.0+0.1	**0.0**
7	0.2+0.2	0.0+0.0	+0.1+0.1	**0.6**
8	0.2+0.2	-0.3-0.3	0.0+0.1	**-0.1**
9	0.0+0.2	-0.3-0.3	0.1+0.0	**-0.3**
10	0.2+0.0	0.0+0.0	0.0+0.1	**0.3**

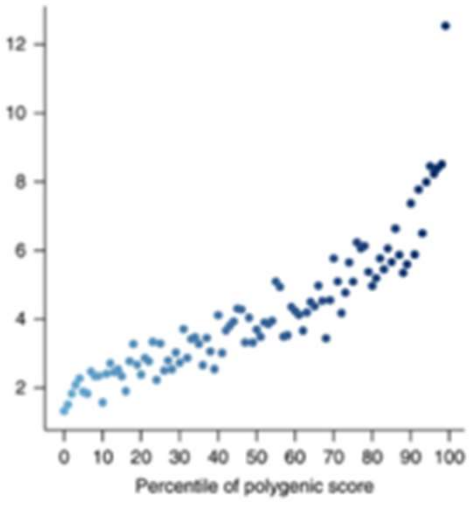
disease risk prediction



Prevalence of Coronary Artery Disease



Prevalence of Type 2 Diabetes

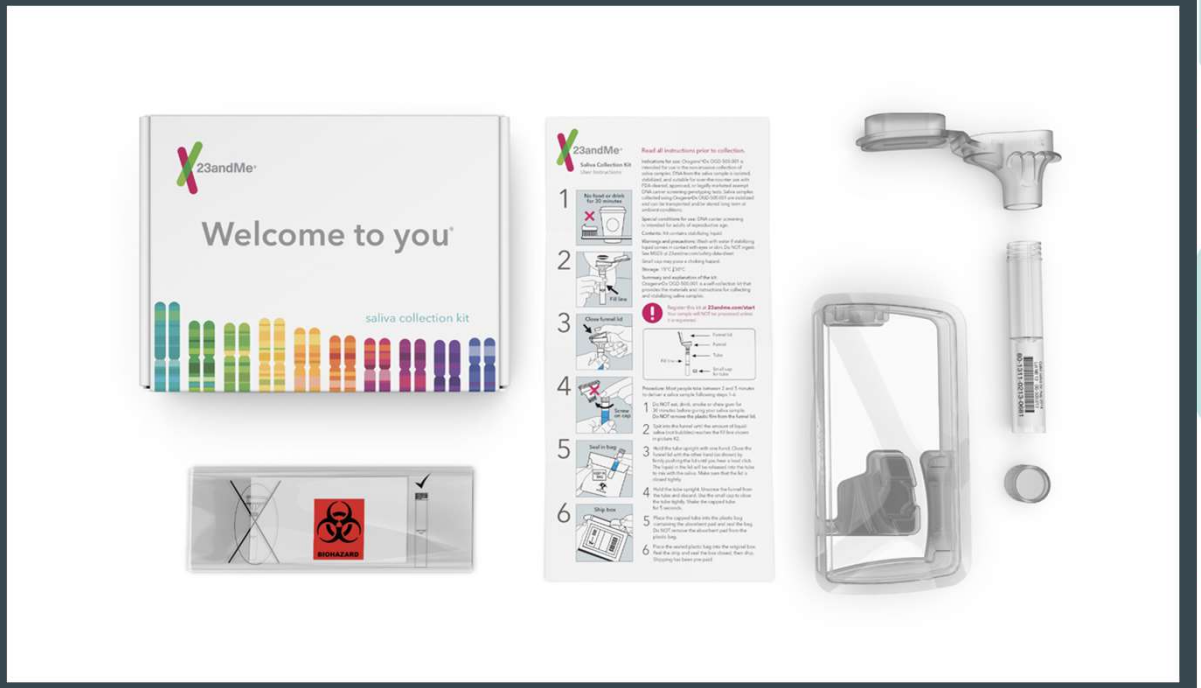


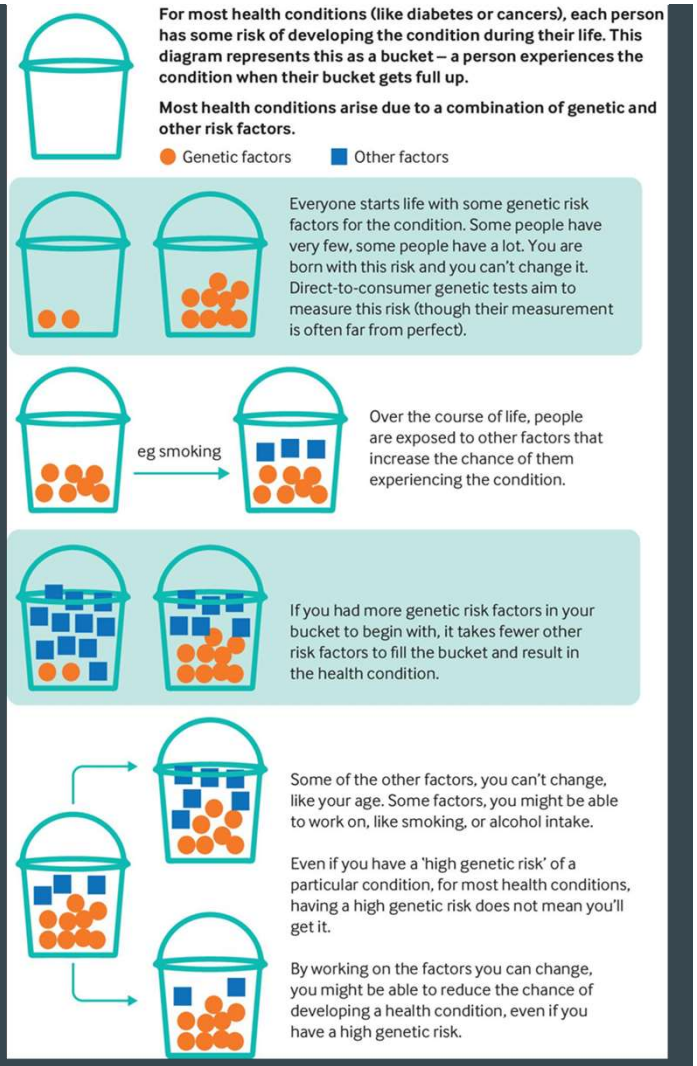
Prevalence of Breast Cancer

"…it is time to contemplate the inclusion of polygenic risk prediction in clinical care, and discuss relevant issues."

Khera et al. Nature Genetics 50, 1219–1224 (2018)

Direct to consumer genetic testing (DTG)





BMJ 2019;367:l5688

1

Taiwan Biobank PheWeb

Browse GWAS results using SAIGE

GSTP

GSTP (alias for **GSTP**1)

GSTP1

2

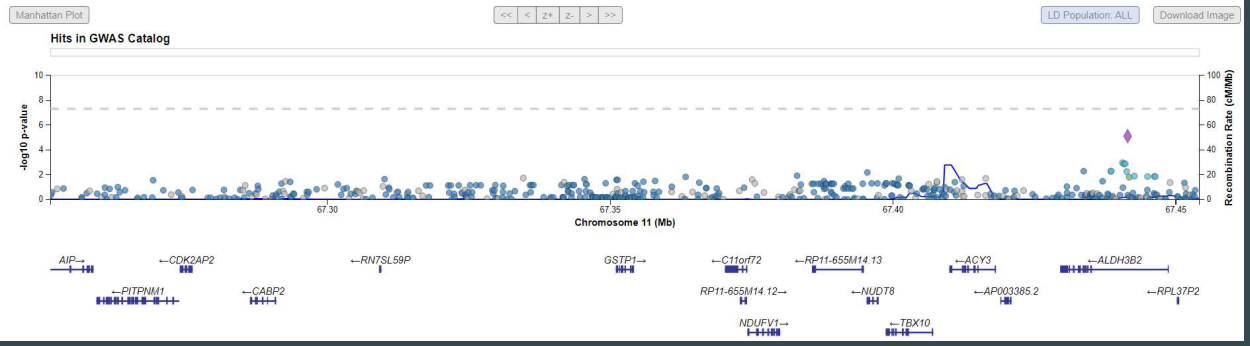
GSTP1

Phenotypes with the most-significant associations for this locus:

Top p-value in gene	Phenotype
7.9e-6	predicted forced expiratory volume in 1 second
1.5e-5	Body+height
1.6e-5	coronary artery disease

64300 samples

Category: **Pulmonary function measurements**



Taiwan Biobank PheWeb

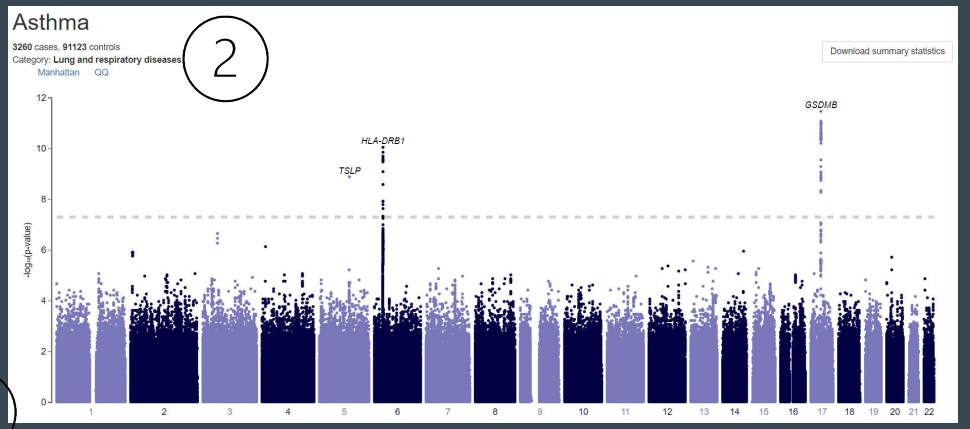
Browse GWAS results using SAIGE

a

- Apoplexia ()
- Arrhythmia ()
- Arthritis ()
- Asthma ()
- Average+personal+income ()
- Drug+allergy ()
- WHR+adjusted+for+BMI (Waist-to-hip ratio adjusted for BMI)

1

2



3

Top Loci:

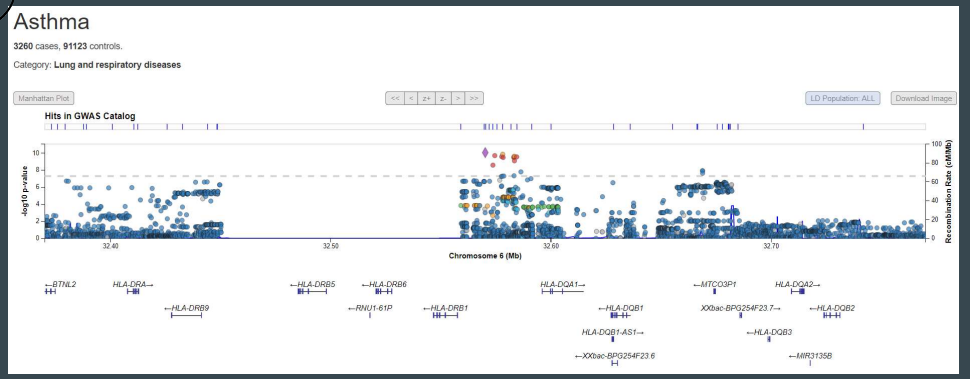
Search... "TCF7L2", "rs1861867", etc.

6 total variants

Variant	Nearest Gene(s)	MAF
17:38,066,267 G / T (rs1008723)	*GSDMB*	0.26
6:32,570,149 A / G (rs2454137)	*HLA-DRB1*	0.37
5:110,401,872 T / C (rs1837253)	*TSLP*	0.37
3:54,464,879 G / T (rs140359708)	*CACNA2D3*	0.019
6:31,850,973 G / A (rs189984590)	*EHMT2*	0.028
4:14,120,444 T / C (rs573242823)	*BOD1L1*	0.0019

← Previous 1 Next →

4



He is the oldest known representation of a being that does not exist in physical form but symbolises ideas about the supernatural.