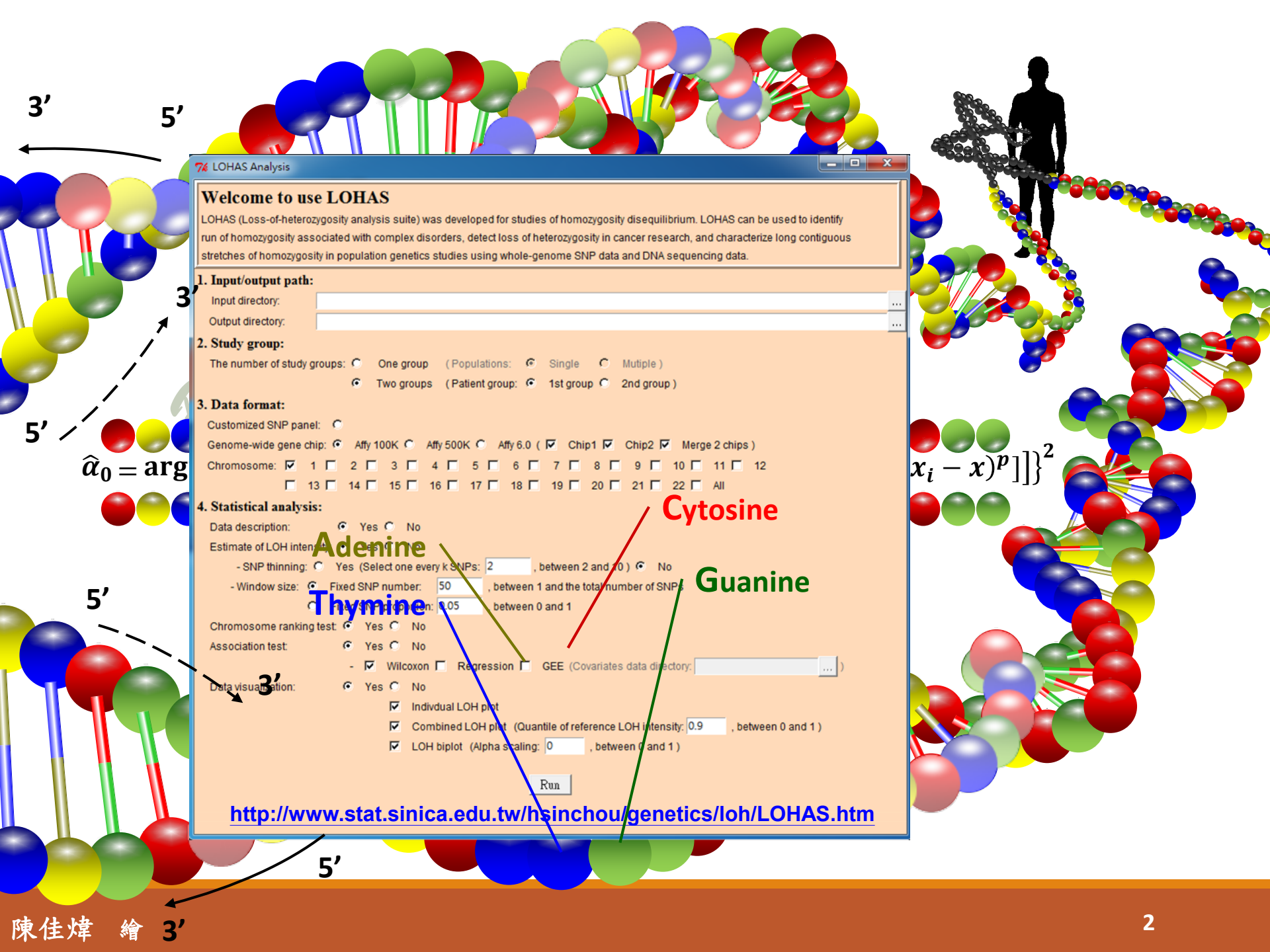
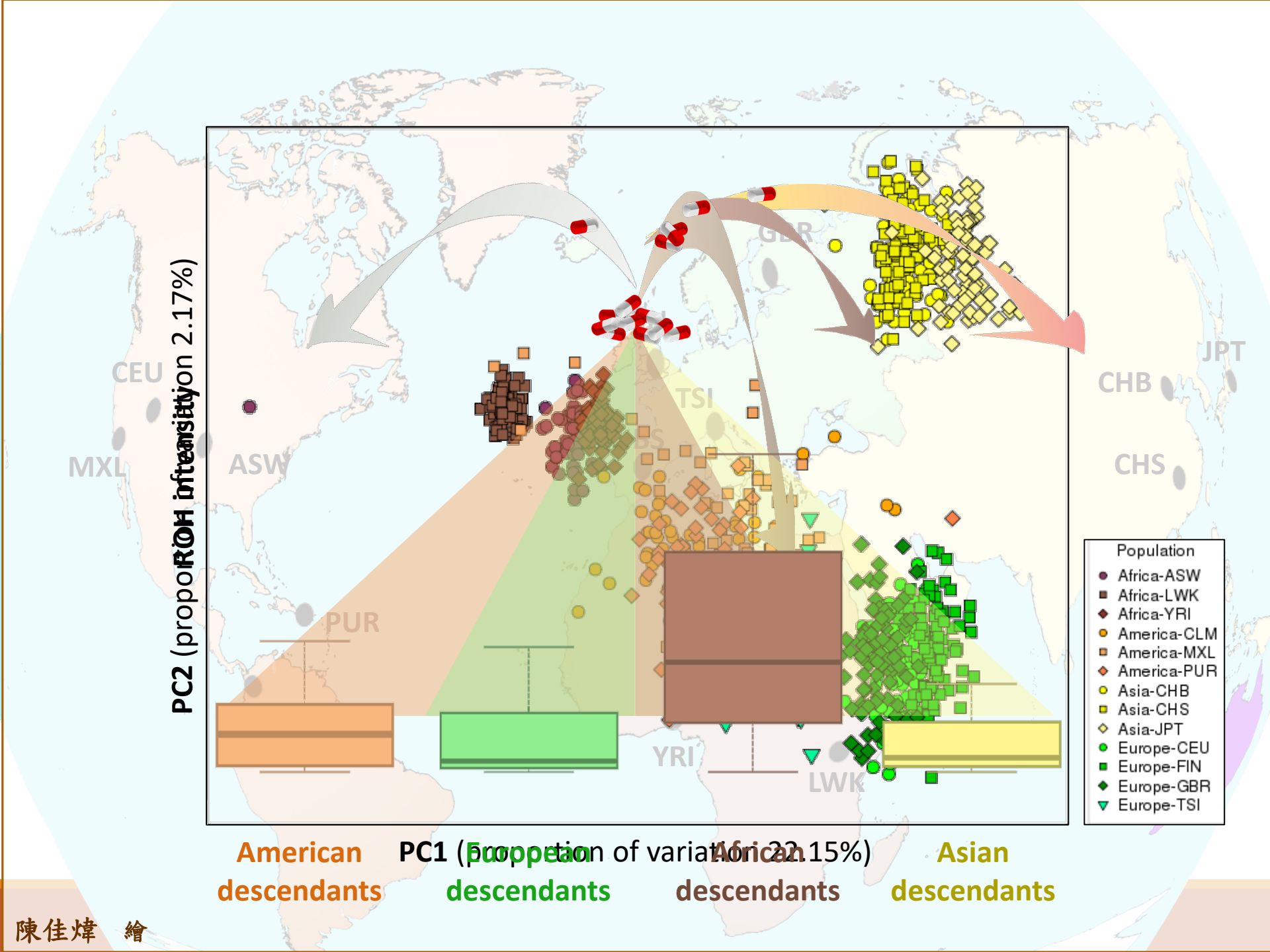
致病基因定位與精準健康 (Disease gene mapping and precision health)

楊 欣 洲 博士 (Hsin-Chou Yang, Ph.D.)

中央研究院 統計科學研究所

電子郵件：hsinchou@stat.sinica.edu.tw
個人網頁：http://www.stat.sinica.edu.tw/hsinchou/
電話：02-2787-5686
傳真：02-2783-1523


3’
5’
3’
5’
5’
3’
5’
3’
$\hat{\alpha}_0 = \arg$
$x_i - x)^p]]\}^2$
LOHAS Analysis
Welcome to use LOHAS
LOHAS (Loss-of-heterozygosity analysis suite) was developed for studies of homozygosity disequilibrium. LOHAS can be used to identify run of homozygosity associated with complex disorders, detect loss of heterozygosity in cancer research, and characterize long contiguous stretches of homozygosity in population genetics studies using whole-genome SNP data and DNA sequencing data.
1. Input/output path:
Input directory:
Output directory:
2. Study group:
The number of study groups: One group (Populations: Single Mutiple)
Two groups (Patient group: 1st group 2nd group)
3. Data format:
Customized SNP panel:
Genome-wide gene chip: Affy 100K Affy 500K Affy 6.0 (Chip1 Chip2 Merge 2 chips)
Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 All
4. Statistical analysis:
Data description: Yes No
Estimate of LOH intensity: Yes
- SNP thinning: Yes (Select one every k SNPs: 2 , between 2 and 10) No
- Window size: Fixed SNP number: 50 , between 1 and the total number of SNPs
between 0 and 1
Chromosome ranking test: Yes No
Association test: Yes No
- Wilcoxon Regression GEE (Covariates data directory:)
Data visualization: Yes No
Indivdual LOH plot
Combined LOH plot (Quantile of reference LOH intensity: 0.9 , between 0 and 1)
LOH biplot (Alpha scaling: 0 , between 0 and 1)
Run
Adenine
Cytosine
Guanine
Thymine
http://www.stat.sinica.edu.tw/hsinchou/genetics/loh/LOHAS.htm

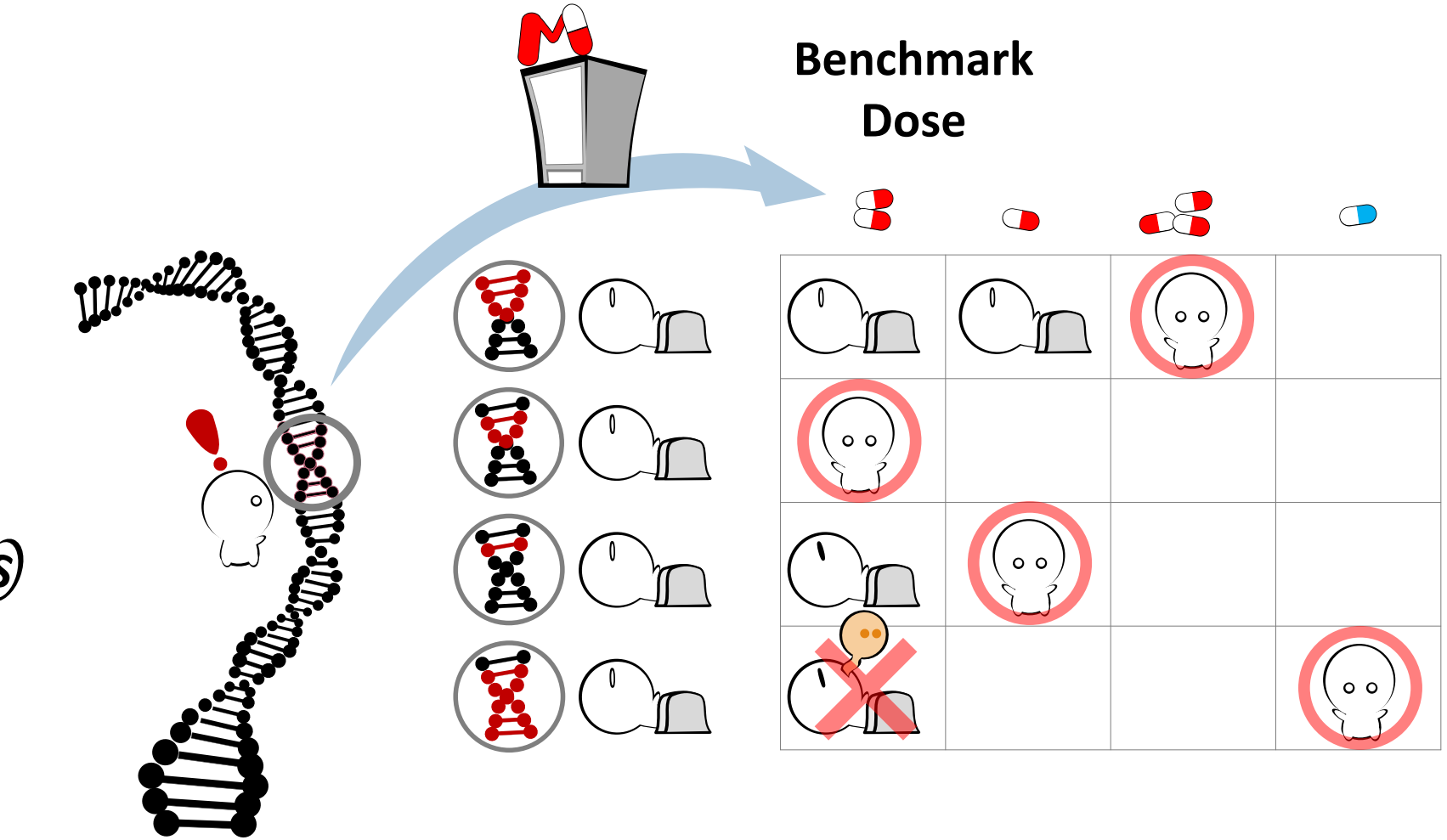

PC2 (proportion of variation 2.17%)
PC1 (proportion of variation 22.15%)
CEU
MXL
ASW
PUR
GBR
TSI
YRI
LWK
JPT
CHB
CHS
Population
Africa-ASW
Africa-LWK
Africa-YRI
America-CLM
America-MXL
America-PUR
Asia-CHB
Asia-CHS
Asia-JPT
Europe-CEU
Europe-FIN
Europe-GBR
Europe-TSI
American descendants
European descendants
African descendants
Asian descendants

陳佳煒 繪

族群與藥物基因體研究

rs1801133 on *MTHFR*

精準健康



個人化醫療與精準醫學

Carbamazepine（卡巴氮平）是一種便宜且止痛效果良好的藥物，常常被拿來用在牙痛、頭痛、癲癇等病症治療，但可能發生嚴重藥物過敏與不良反應（亦即，史蒂芬強生症候群 (Steven-Johnson Syndrome, SJS)），因為此藥通報不良反應之死亡案例高達四例！

Table 1 **Frequency of *HLA* alleles in patients with Stevens–Johnson syndrome**

HLA allele	*CBZ–SJS*	*CBZ-tolerant*	*Normal*
*B*1502*	44 (100%)	3 (3%)*	8 (8.6%)†
*Cw*0801*	41 (93.2%)	17 (16.8%)	13 (14%)
*A*1101*	36 (81.8%)	51 (50.5%)	53 (57%)
*DRB1*1202*	33 (75%)	12 (11.9%)	18 (19.4%)
*B*1502, Cw*0801*	41 (93.2%)	3 (3%)	7 (7.5%)
*B*1502, A*1101*	36 (81.8%)	2 (2%)	6 (6.5%)
*B*1502, DRB1*1202*	33 (75%)	1(1%)	5 (5.4%)
*B*1502, Cw*0801, A*1101, DRB1*1202*	29(66%)	0 (0%)	3 (3.2%)

Frequencies (by number and percentage) of individual or combined loci of the *B*1502* ancestral haplotype are shown in patients with carbamazepine-induced Stevens–Johnson syndrome (CBZ–SJS; $n = 44$), and in carbamazepine-tolerant ($n = 101$) and normal subjects ($n = 93$). For methods, see supplementary information.

*Odds ratio (CBZ–SJS/CBZ-tolerant): 2,504 (95% CI, 126–49,522); corrected *P* value $P_c = 3.13 \times 10^{-27}$.

†Odds ratio (CBZ–SJS/normal): 895 (95% CI, 50–15,869); $P_c = 1.38 \times 10^{-21}$.

健保給付之基因檢測

檢驗細項　　表單編號: EDCP-QP-1214/01

檢驗項目	HLA-B 1502(for Carbamazepine)	院內醫令碼	L1219601
中文名稱	HLA-B 1502 基因檢測	健保代碼	12196B
適用檢體別	血液	支付點數	2648
建議採檢容器	紫頭管【K_2EDTA】	檢驗操作方法	Real-time PCR
檢體採集量	3mL		
送檢方式	人工傳送/氣送	參考值(單位)	Non-detected
收檢時間	24小時		
報告時效	7天	採檢容器料號	M3110016
操作組別/分機	基因診斷組 / 5807	其他備註事項	
檢驗操作時間	週一至週五 08:00 - 17:30		
採檢應注意事項	採檢後，請180度輕微搖晃4~5次，以利抗凝劑混合均勻。		

臨床意義與用途：

HLA基因位在人類的第六條染色體的短臂上，共分成兩大類，第一類有A、B、C三種，第二類有DP、DQ、DR；不同人類個體之間的HLA具有高度多型性(Highly polymorphism)，各自的HLA型別是來自己的父母親，所以兄弟姐妹之間有接近四分之一的機會完全相同，若父親或母親有一位是HLA-B*1502的帶因者，則該夫妻的小孩也會有1/2的機會是HLA-B*1502的帶因者。

HLA-B*1502基因的存在被證實與漢人族群(Han Chinese)因服用卡巴氮平(Carbamazepine)或是aromatic antiepileptic drugs (AEDs)此類的藥物而引起的史蒂芬強生症候群(Steven-Johnson Syndrome, SJS)，以及臨床上致死率高達30%的毒性上皮溶解症(Toxic Epidermal necrolysis, TEN)有高度的關聯性，臨床上癲癇或是有神經疼痛的病人，在服用卡巴氮平(Carbamazepine)或是AEDs此類的藥物之前，臨床醫師應當考慮先進行HLA-B*1502基因檢測，以加強用藥安全。

註1. 健保每人限申報一次，適應症：癲癇症、三叉神經痛、腎原性尿崩症及雙極性之精神疾患初次需使用含卡 巴氮平Carbamazepine成份藥物之病患。

註2. 本檢驗試劑使用Pharm iGene PG1502 detection Kit，以 即時聚合酶連鎖反應定性檢測HLA-B*1502基因。根據原廠說明，此檢驗相對於基因定序方法，其特異性可達100%，敏感度可達99.78%，偵測極限(detection limit)為25ng/reaction；無法區分以下罕見的基因型別:HLA-B*1513,1531,1555,1588, 1589,1820,9512,9521,9544及9570。

註3. 本試驗偵測極限(detection limit)為10 ng/reaction，因此檢測檢體DNA濃度應維持於5.0 ng/ul，方能維持檢驗品質，濃度不足時將於報告中備註。

楔子：

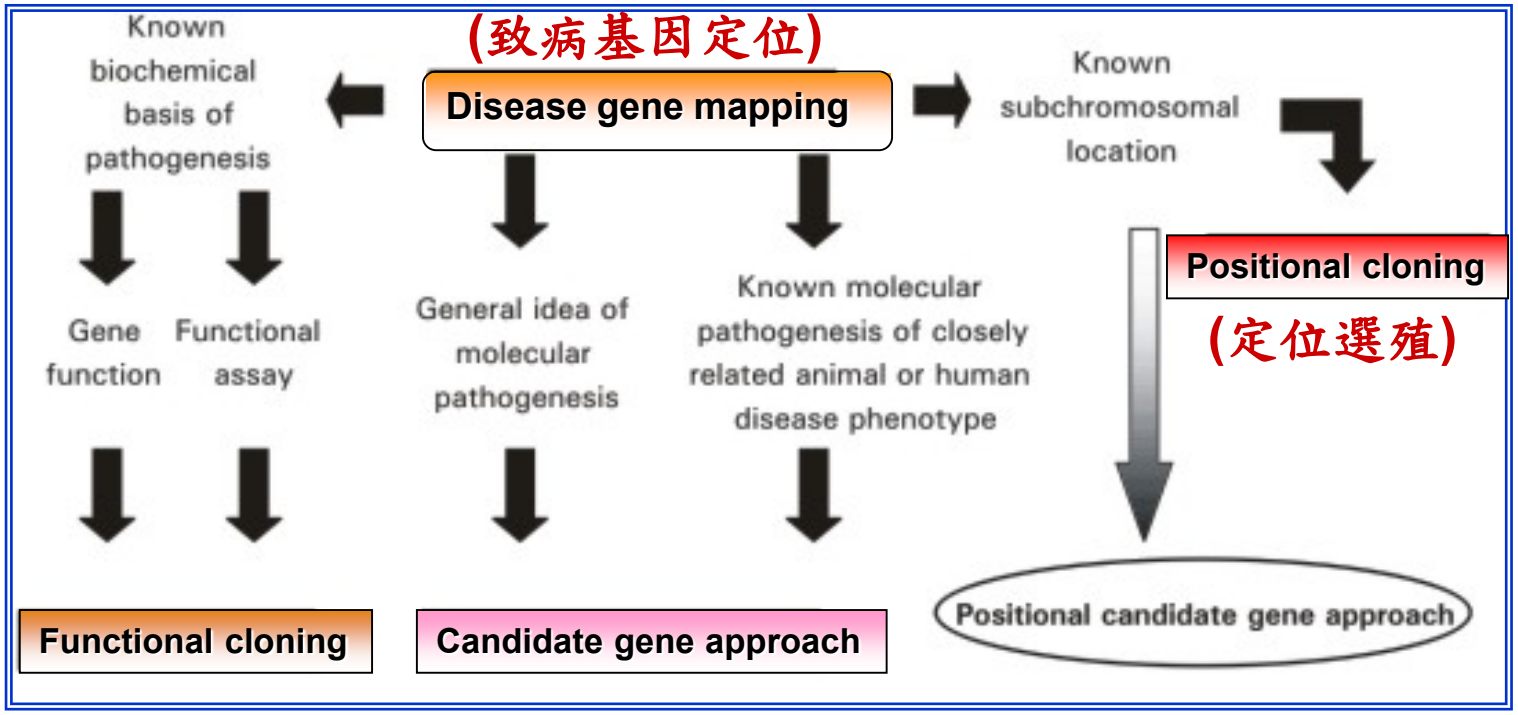
要戰勝敵人（疾病），必須先找出敵人（致病基因）

兩個抓壞蛋的故事

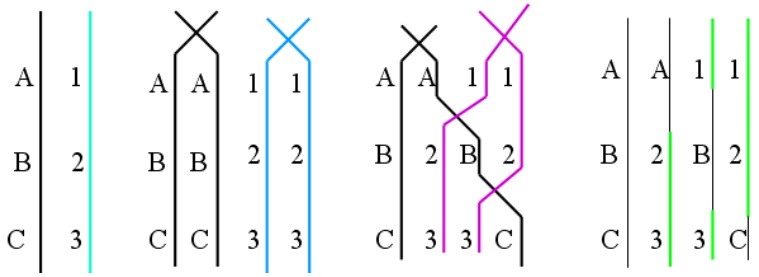
故事	情節一	情節二
主軸內容	警察捉搶匪	致病基因定位
故事背景	搶匪擾亂社會治安	基因變異危害人體健康
主人翁	英勇的警方	強力的統計檢定
主要任務	捉搶匪	找致病基因
藏匿範圍	台澎金馬	人類基因體
責任分區	二十二縣市	二十三對染色體
受檢對象	兩千三百萬居民	三十億對的鹼基對
追捕工具	無線電收發器	標誌基因

致病基因定位的策略

致病基因定位：找到致病基因在染色體上的位置的程序



名詞



Chromosome（染色體）is a long string double-stranded helix DNA. **Locus**（位點） denotes a position on a chromosome. **Marker**（標誌基因）denotes a locus-specific polymorphism. **Allele**（等位基因 or 對偶基因）is a type of DNA at a particular locus. **Haplotype**（單套型）is a set of alleles of closely linked genetic markers present on one chromosome. **Genotype**（基因型）denotes a pair of alleles at a particular locus. **Phenotype**（外表型）is a characteristic of a subject.

標誌基因

大標記：短片段重複序列

小標記：單一核苷酸

短片段重複序列多型性（STRP）

D7S280

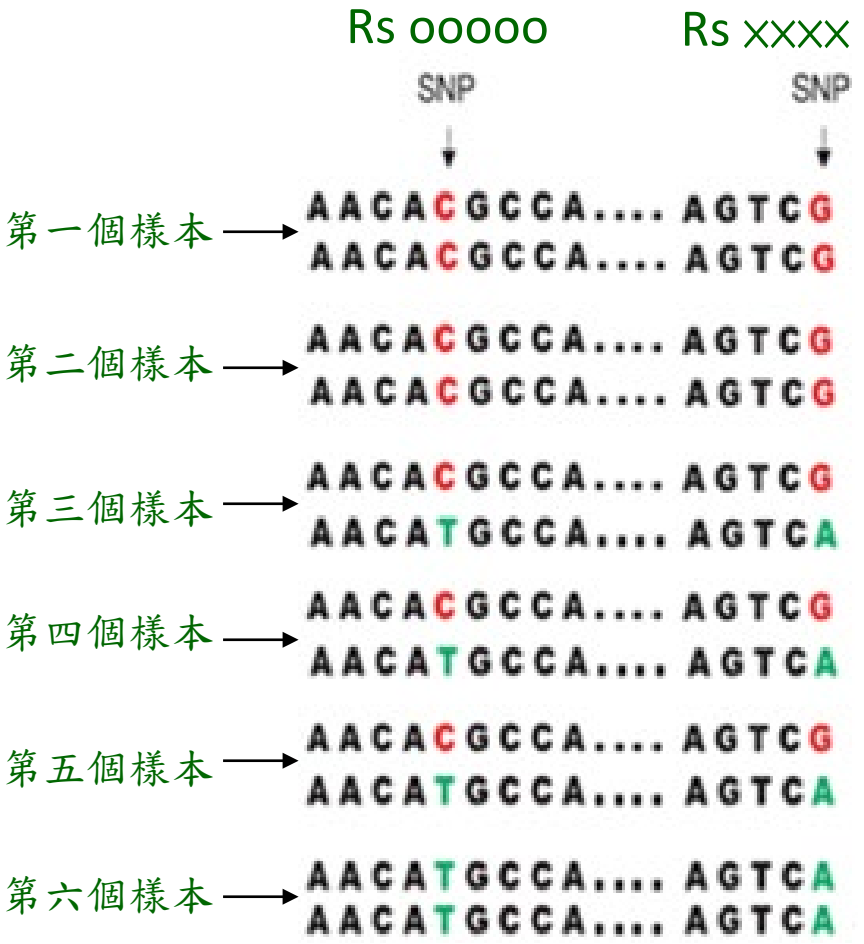
```
  1 aatttttgta ttttttttag agacggggtt tcaccatgtt ggtcaggctg actatggagt
 61 tattttaagg ttaatatata taaagggtat gatagaacac ttgtcatagt ttagaacgaa
121 ctaacgatag atagatagat agatagatag atagatagat agatagatag atagacagat
181 tgatagtttt tttttatctc actaaatagt ctatagtaaa catttaatta ccaatatttg
241 gtgcaattct gtcaatgagg ataaatgtgg aatcgttata attcttaaga atatatattc
301 cctctgagtt tttgatacct cagattttaa ggcc

  1 aatttttgta ttttttttag agacggggtt tcaccatgtt ggtcaggctg actatggagt
 61 tattttaagg ttaatatata taaagggtat gatagaacac ttgtcatagt ttagaacgaa
121 ctaacgatag atagatagat agatagatag atagatagat agatagatag atagacagat
181 tgatagatag atagatagat agatagatag ttttttttta tctcactaaa tagtctatag
241 taaacattta attaccaata tttggtgcaa ttctgtcaat gaggataaat gtggaatcgt
301 tataattctt aagaatatat attccctctg agtt
```

(gata) (15, 20)

毛細管電泳鑑定

單一核苷酸多型性（SNP）

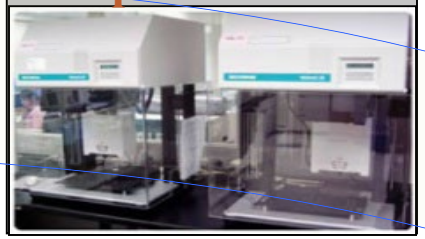


MALDI-TOF 質譜儀

PCR

PCR Farm

SpectroPREP

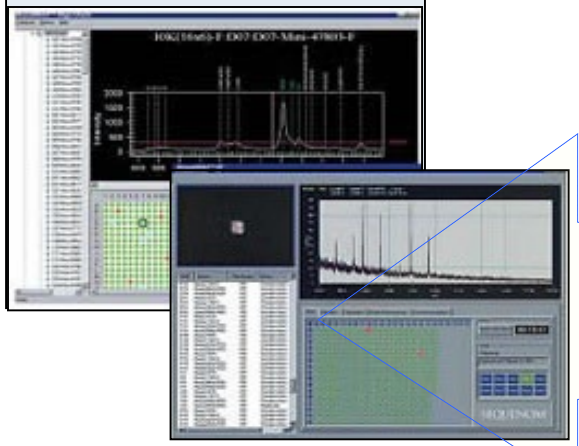


液體分注工作站

SpectroPOINT

MALDI Matrix Array點製機

MassARRAY

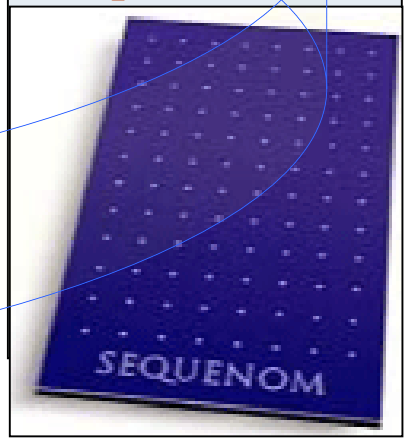


包含assay 及primer設計軟體、Oligo Check軟體、即時質譜分析基因型分析軟體、以及後端整合資料庫

SpectroTYPER

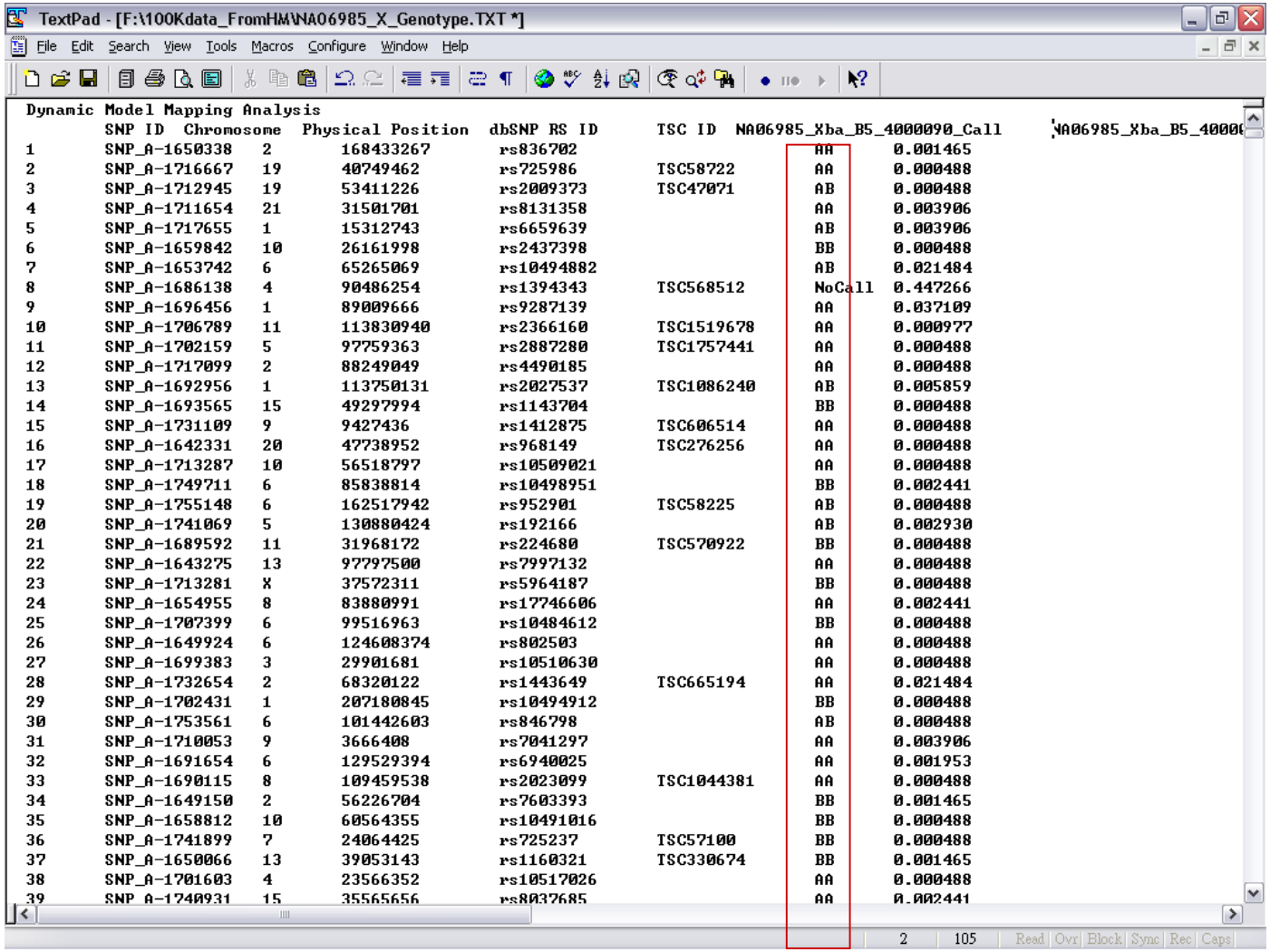
MALDI-TOF 質譜分析儀

SpectroHIP



MALDI matrix晶片

SNP基因型資料（genotype data）


TextPad - [F:\100Kdata_FromHM\NA06985_X_Genotype.TXT *]

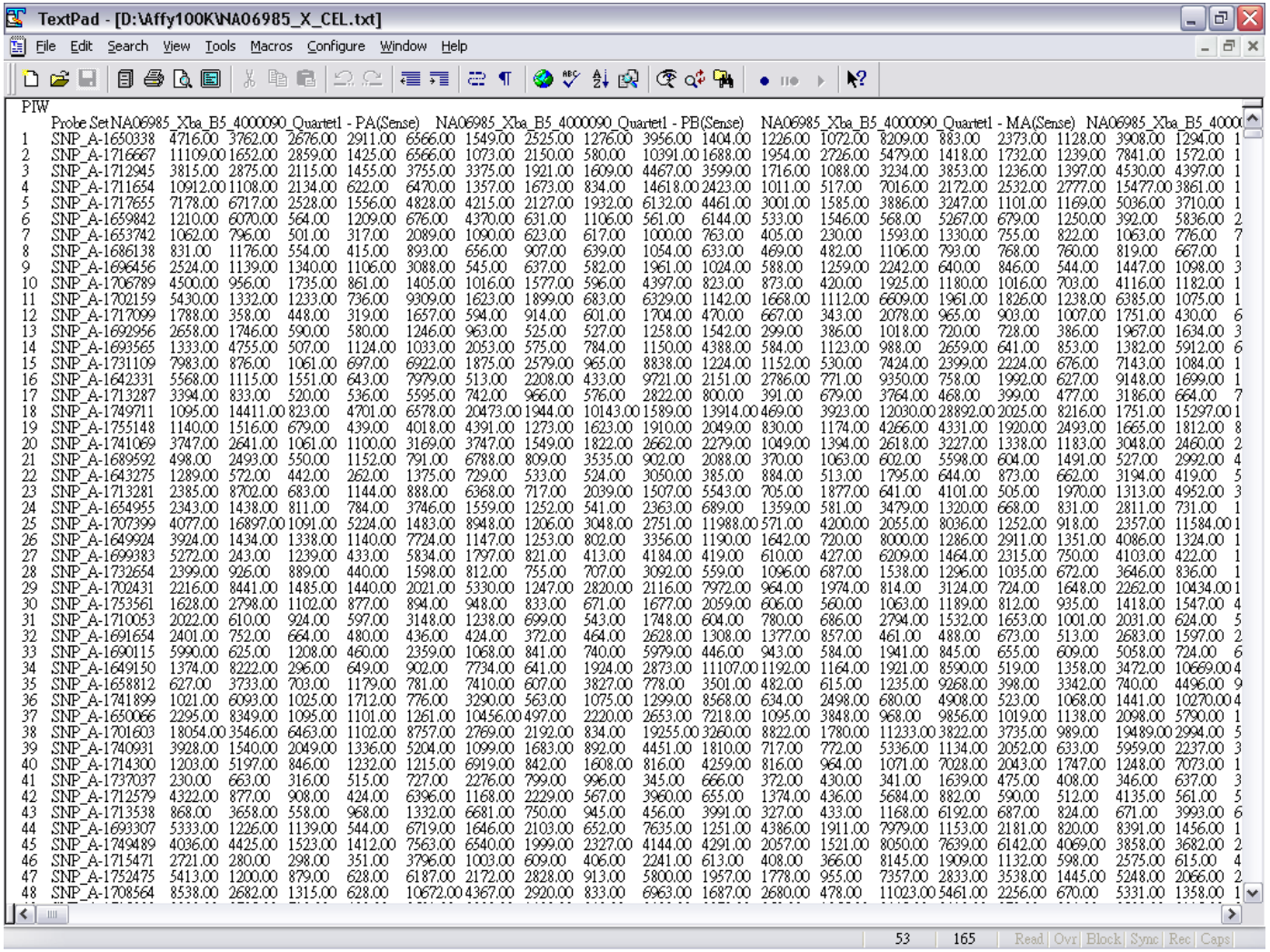
File Edit Search View Tools Macros Configure Window Help

Dynamic Model Mapping Analysis

	SNP ID	Chromosome	Physical Position	dbSNP RS ID	TSC ID	NA06985_Xba_B5_4000090_Call	NA06985_Xba_B5_4000
1	SNP_A-1650338	2	168433267	rs836702		AA	0.001465
2	SNP_A-1716667	19	40749462	rs725986	TSC58722	AA	0.000488
3	SNP_A-1712945	19	53411226	rs2009373	TSC47071	AB	0.000488
4	SNP_A-1711654	21	31501701	rs8131358		AA	0.003906
5	SNP_A-1717655	1	15312743	rs6659639		AB	0.003906
6	SNP_A-1659842	10	26161998	rs2437398		BB	0.000488
7	SNP_A-1653742	6	65265069	rs10494882		AB	0.021484
8	SNP_A-1686138	4	90486254	rs1394343	TSC568512	NoCall	0.447266
9	SNP_A-1696456	1	89009666	rs9287139		AA	0.037109
10	SNP_A-1706789	11	113830940	rs2366160	TSC1519678	AA	0.000977
11	SNP_A-1702159	5	97759363	rs2887280	TSC1757441	AA	0.000488
12	SNP_A-1717099	2	88249049	rs4490185		AA	0.000488
13	SNP_A-1692956	1	113750131	rs2027537	TSC1086240	AB	0.005859
14	SNP_A-1693565	15	49297994	rs1143704		BB	0.000488
15	SNP_A-1731109	9	9427436	rs1412875	TSC606514	AA	0.000488
16	SNP_A-1642331	20	47738952	rs968149	TSC276256	AA	0.000488
17	SNP_A-1713287	10	56518797	rs10509021		AA	0.000488
18	SNP_A-1749711	6	85838814	rs10498951		BB	0.002441
19	SNP_A-1755148	6	162517942	rs952901	TSC58225	AB	0.000488
20	SNP_A-1741069	5	130880424	rs192166		AB	0.002930
21	SNP_A-1689592	11	31968172	rs224680	TSC570922	BB	0.000488
22	SNP_A-1643275	13	97797500	rs7997132		AA	0.000488
23	SNP_A-1713281	X	37572311	rs5964187		BB	0.000488
24	SNP_A-1654955	8	83880991	rs17746606		AA	0.002441
25	SNP_A-1707399	6	99516963	rs10484612		BB	0.000488
26	SNP_A-1649924	6	124608374	rs802503		AA	0.000488
27	SNP_A-1699383	3	29901681	rs10510630		AA	0.000488
28	SNP_A-1732654	2	68320122	rs1443649	TSC665194	AA	0.021484
29	SNP_A-1702431	1	207180845	rs10494912		BB	0.000488
30	SNP_A-1753561	6	101442603	rs846798		AB	0.000488
31	SNP_A-1710053	9	3666408	rs7041297		AA	0.003906
32	SNP_A-1691654	6	129529394	rs6940025		AA	0.001953
33	SNP_A-1690115	8	109459538	rs2023099	TSC1044381	AA	0.000488
34	SNP_A-1649150	2	56226704	rs7603393		BB	0.001465
35	SNP_A-1658812	10	60564355	rs10491016		BB	0.000488
36	SNP_A-1741899	7	24064425	rs725237	TSC57100	BB	0.000488
37	SNP_A-1650066	13	39053143	rs1160321	TSC330674	BB	0.001465
38	SNP_A-1701603	4	23566352	rs10517026		AA	0.000488
39	SNP_A-1740931	15	35565656	rs8037685		AA	0.002441

2 105 Read Ovr Block Sync Rec Caps

雜交強度資料（Hybridization intensity）



TextPad - [D:\Affy100K\NA06985_X_CEL.txt]

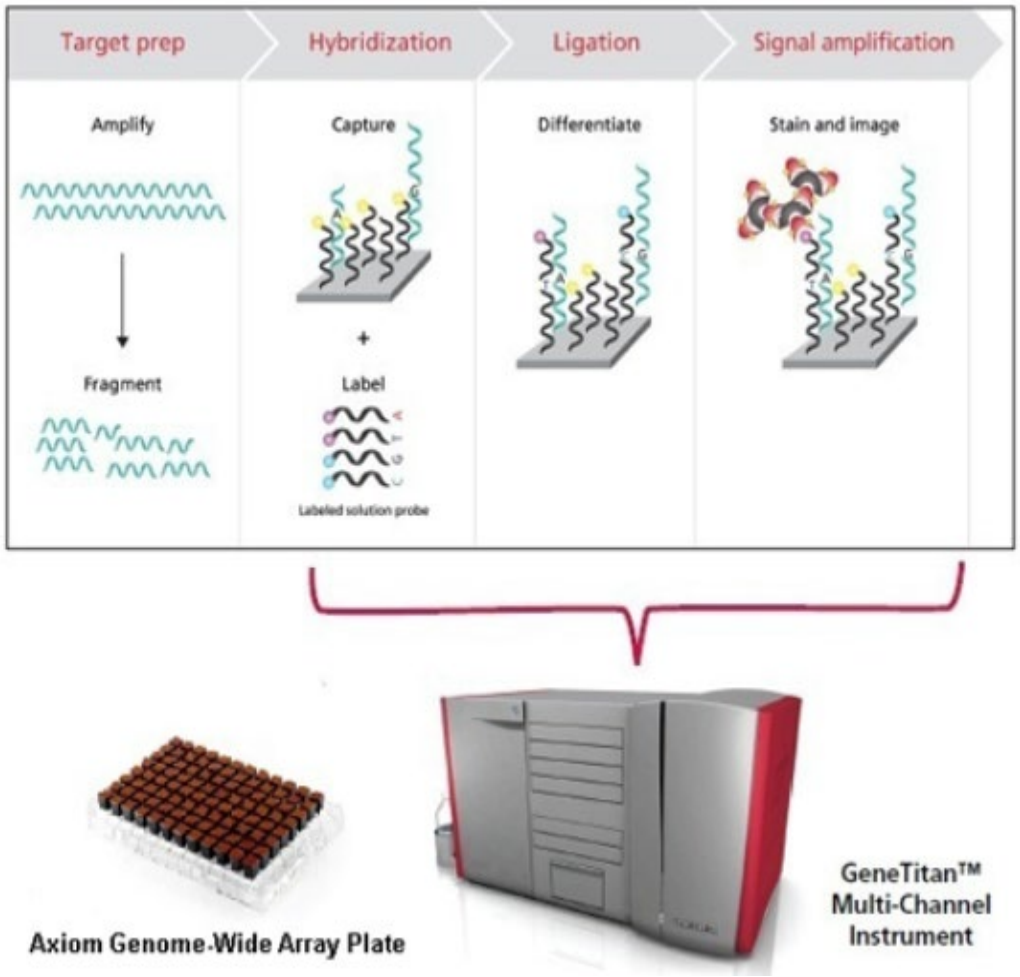
File Edit Search View Tools Macros Configure Window Help

PIW

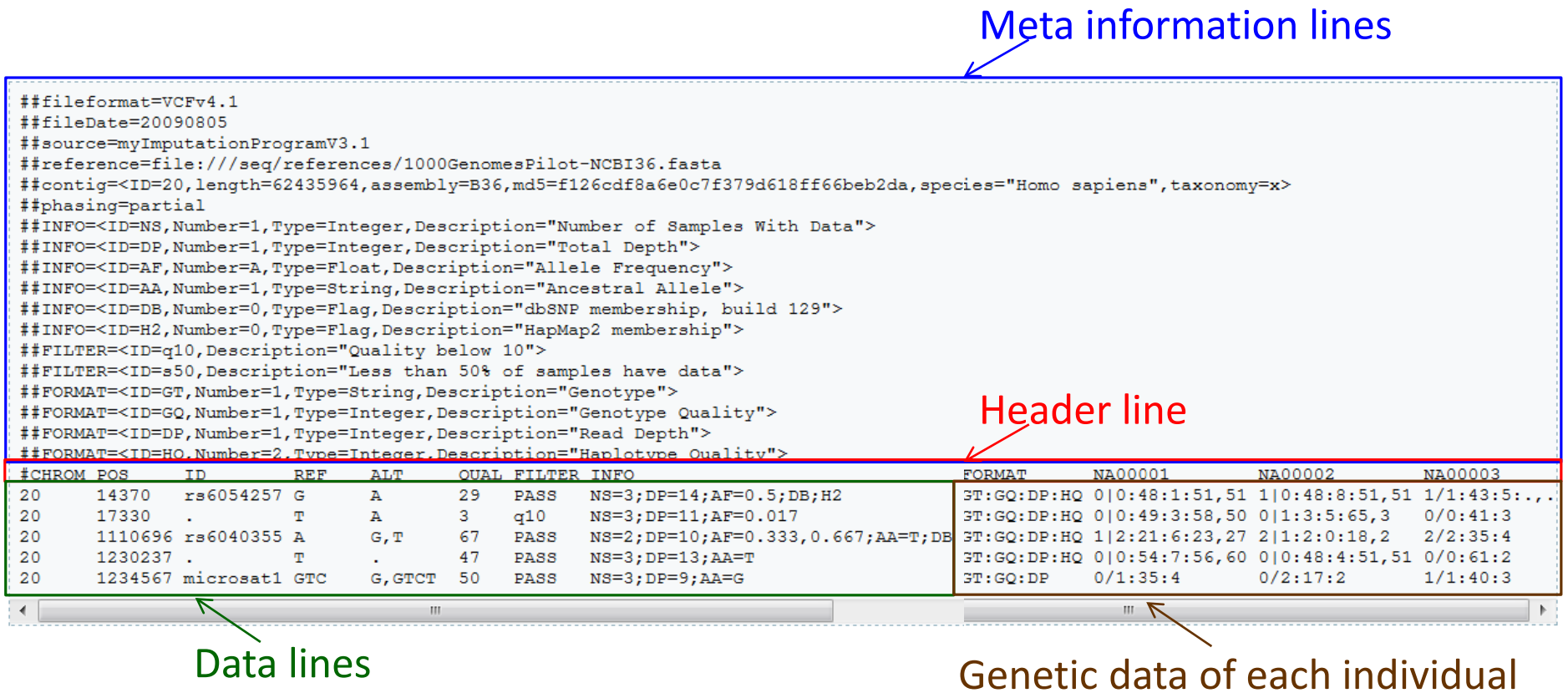
	Probe Set	NA06985_Xba_B5_4000090_Quartet1 - PA(Sense)					NA06985_Xba_B5_4000090_Quartet1 - PB(Sense)					NA06985_Xba_B5_4000090_Quartet1 - MA(Sense)					NA06985_Xba_B5_4000	
1	SNP_A-1650338	4716.00	3762.00	2676.00	2911.00	6566.00	1549.00	2525.00	1276.00	3956.00	1404.00	1226.00	1072.00	8209.00	883.00	2373.00	1128.00	3908.00	1294.00	1
2	SNP_A-1716667	11109.00	1652.00	2859.00	1425.00	6566.00	1073.00	2150.00	580.00	10391.00	1688.00	1954.00	2726.00	5479.00	1418.00	1732.00	1239.00	7841.00	1572.00	1
3	SNP_A-1712945	3815.00	2875.00	2115.00	1455.00	3755.00	3375.00	1921.00	1609.00	4467.00	3599.00	1716.00	1088.00	3234.00	3853.00	1236.00	1397.00	4530.00	4397.00	1
4	SNP_A-1711654	10912.00	1108.00	2134.00	622.00	6470.00	1357.00	1673.00	834.00	14618.00	2423.00	1011.00	517.00	7016.00	2172.00	2532.00	2777.00	15477.00	3861.00	1
5	SNP_A-1717655	7178.00	6717.00	2528.00	1556.00	4828.00	4215.00	2127.00	1932.00	6132.00	4461.00	3001.00	1585.00	3886.00	3247.00	1101.00	1169.00	5036.00	3710.00	1
6	SNP_A-1659842	1210.00	6070.00	564.00	1209.00	676.00	4370.00	631.00	1106.00	561.00	6144.00	533.00	1546.00	568.00	5267.00	679.00	1250.00	392.00	5836.00	2
7	SNP_A-1653742	1062.00	796.00	501.00	317.00	2089.00	1090.00	623.00	617.00	1000.00	763.00	405.00	230.00	1593.00	1330.00	755.00	822.00	1063.00	776.00	7
8	SNP_A-1686138	831.00	1176.00	554.00	415.00	893.00	656.00	907.00	639.00	1054.00	633.00	469.00	482.00	1106.00	793.00	768.00	760.00	819.00	667.00	1
9	SNP_A-1696456	2524.00	1139.00	1340.00	1106.00	3088.00	545.00	637.00	582.00	1961.00	1024.00	588.00	1259.00	2242.00	640.00	846.00	544.00	1447.00	1098.00	3
10	SNP_A-1706789	4500.00	956.00	1735.00	861.00	1405.00	1016.00	1577.00	596.00	4397.00	823.00	873.00	420.00	1925.00	1180.00	1016.00	703.00	4116.00	1182.00	1
11	SNP_A-1702159	5430.00	1332.00	1233.00	736.00	9309.00	1623.00	1899.00	683.00	6329.00	1142.00	1668.00	1112.00	6609.00	1961.00	1826.00	1238.00	6385.00	1075.00	1
12	SNP_A-1717099	1788.00	358.00	448.00	319.00	1657.00	594.00	914.00	601.00	1704.00	470.00	667.00	343.00	2078.00	965.00	903.00	1007.00	1751.00	430.00	6
13	SNP_A-1692956	2658.00	1746.00	590.00	580.00	1246.00	963.00	525.00	527.00	1258.00	1542.00	299.00	386.00	1018.00	720.00	728.00	386.00	1967.00	1634.00	3
14	SNP_A-1693565	1333.00	4755.00	507.00	1124.00	1033.00	2053.00	575.00	784.00	1150.00	4388.00	584.00	1123.00	988.00	2659.00	641.00	853.00	1382.00	5912.00	6
15	SNP_A-1731109	7983.00	876.00	1061.00	697.00	6922.00	1875.00	2579.00	965.00	8838.00	1224.00	1152.00	530.00	7424.00	2399.00	2224.00	676.00	7143.00	1084.00	1
16	SNP_A-1642331	5568.00	1115.00	1551.00	643.00	7979.00	513.00	2208.00	433.00	9721.00	2151.00	2786.00	771.00	9350.00	758.00	1992.00	627.00	9148.00	1699.00	1
17	SNP_A-1713287	3394.00	833.00	520.00	536.00	5595.00	742.00	966.00	576.00	2822.00	800.00	391.00	679.00	3764.00	468.00	399.00	477.00	3186.00	664.00	7
18	SNP_A-1749711	1095.00	14411.00	823.00	4701.00	6578.00	20473.00	1944.00	10143.00	1589.00	13914.00	469.00	3923.00	12030.00	28892.00	2025.00	8216.00	1751.00	15297.00	1
19	SNP_A-1755148	1140.00	1516.00	679.00	439.00	4018.00	4391.00	1273.00	1623.00	1910.00	2049.00	830.00	1174.00	4266.00	4331.00	1920.00	2493.00	1665.00	1812.00	8
20	SNP_A-1741069	3747.00	2641.00	1061.00	1100.00	3169.00	3747.00	1549.00	1822.00	2662.00	2279.00	1049.00	1394.00	2618.00	3227.00	1338.00	1183.00	3048.00	2460.00	2
21	SNP_A-1689592	498.00	2493.00	550.00	1152.00	791.00	6788.00	809.00	3535.00	902.00	2088.00	370.00	1063.00	602.00	5598.00	604.00	1491.00	527.00	2992.00	4
22	SNP_A-1643275	1289.00	572.00	442.00	262.00	1375.00	729.00	533.00	524.00	3050.00	385.00	884.00	513.00	1795.00	644.00	873.00	662.00	3194.00	419.00	5
23	SNP_A-1713281	2385.00	8702.00	683.00	1144.00	888.00	6368.00	717.00	2039.00	1507.00	5543.00	705.00	1877.00	641.00	4101.00	505.00	1970.00	1313.00	4952.00	3
24	SNP_A-1654955	2343.00	1438.00	811.00	784.00	3746.00	1559.00	1252.00	541.00	2363.00	689.00	1359.00	581.00	3479.00	1320.00	668.00	831.00	2811.00	731.00	1
25	SNP_A-1707399	4077.00	16897.00	1091.00	5224.00	1483.00	8948.00	1206.00	3048.00	2751.00	11988.00	571.00	4200.00	2055.00	8036.00	1252.00	918.00	2357.00	11584.00	1
26	SNP_A-1649924	3924.00	1434.00	1338.00	1140.00	7724.00	1147.00	1253.00	802.00	3356.00	1190.00	1642.00	720.00	8000.00	1286.00	2911.00	1351.00	4086.00	1324.00	1
27	SNP_A-1699383	5272.00	243.00	1239.00	433.00	5834.00	1797.00	821.00	413.00	4184.00	419.00	610.00	427.00	6209.00	1464.00	2315.00	750.00	4103.00	422.00	1
28	SNP_A-1732654	2399.00	926.00	889.00	440.00	1598.00	812.00	755.00	707.00	3092.00	559.00	1096.00	687.00	1538.00	1296.00	1035.00	672.00	3646.00	836.00	1
29	SNP_A-1702431	2216.00	8441.00	1485.00	1440.00	2021.00	5330.00	1247.00	2820.00	2116.00	7972.00	964.00	1974.00	814.00	3124.00	724.00	1648.00	2262.00	10434.00	1
30	SNP_A-1753561	1628.00	2798.00	1102.00	877.00	894.00	948.00	833.00	671.00	1677.00	2059.00	606.00	560.00	1063.00	1189.00	812.00	935.00	1418.00	1547.00	4
31	SNP_A-1710053	2022.00	610.00	924.00	597.00	3148.00	1238.00	699.00	543.00	1748.00	604.00	780.00	686.00	2794.00	1532.00	1653.00	1001.00	2031.00	624.00	5
32	SNP_A-1691654	2401.00	752.00	664.00	480.00	436.00	424.00	372.00	464.00	2628.00	1308.00	1377.00	857.00	461.00	488.00	673.00	513.00	2683.00	1597.00	2
33	SNP_A-1690115	5990.00	625.00	1208.00	460.00	2359.00	1068.00	841.00	740.00	5979.00	446.00	943.00	584.00	1941.00	845.00	655.00	609.00	5058.00	724.00	6
34	SNP_A-1649150	1374.00	8222.00	296.00	649.00	902.00	7734.00	641.00	1924.00	2873.00	11107.00	1192.00	1164.00	1921.00	8590.00	519.00	1358.00	3472.00	10669.00	4
35	SNP_A-1658812	627.00	3733.00	703.00	1179.00	781.00	7410.00	607.00	3827.00	778.00	3501.00	482.00	615.00	1235.00	9268.00	398.00	3342.00	740.00	4496.00	9
36	SNP_A-1741899	1021.00	6093.00	1025.00	1712.00	776.00	3290.00	563.00	1075.00	1299.00	8568.00	634.00	2498.00	680.00	4908.00	523.00	1068.00	1441.00	10270.00	4
37	SNP_A-1650066	2295.00	8349.00	1095.00	1101.00	1261.00	10456.00	497.00	2220.00	2653.00	7218.00	1095.00	3848.00	968.00	9856.00	1019.00	1138.00	2098.00	5790.00	1
38	SNP_A-1701603	18054.00	3546.00	6463.00	1102.00	8757.00	2769.00	2192.00	834.00	19255.00	3260.00	8822.00	1780.00	11233.00	3822.00	3735.00	989.00	19489.00	2994.00	5
39	SNP_A-1740931	3928.00	1540.00	2049.00	1336.00	5204.00	1099.00	1683.00	892.00	4451.00	1810.00	717.00	772.00	5336.00	1134.00	2052.00	633.00	5959.00	2237.00	3
40	SNP_A-1714300	1203.00	5197.00	846.00	1232.00	1215.00	6919.00	842.00	1608.00	816.00	4259.00	816.00	964.00	1071.00	7028.00	2043.00	1747.00	1248.00	7073.00	1
41	SNP_A-1737037	230.00	663.00	316.00	515.00	727.00	2276.00	799.00	996.00	345.00	666.00	372.00	430.00	341.00	1639.00	475.00	408.00	346.00	637.00	3
42	SNP_A-1712579	4322.00	877.00	908.00	424.00	6396.00	1168.00	2229.00	567.00	3960.00	655.00	1374.00	436.00	5684.00	882.00	590.00	512.00	4135.00	561.00	5
43	SNP_A-1713538	868.00	3658.00	558.00	968.00	1332.00	6681.00	750.00	945.00	456.00	3991.00	327.00	433.00	1168.00	6192.00	687.00	824.00	671.00	3993.00	6
44	SNP_A-1693307	5333.00	1226.00	1139.00	544.00	6719.00	1646.00	2103.00	652.00	7635.00	1251.00	4386.00	1911.00	7979.00	1153.00	2181.00	820.00	8391.00	1456.00	1
45	SNP_A-1749489	4036.00	4425.00	1523.00	1412.00	7563.00	6540.00	1999.00	2327.00	4144.00	4291.00	2057.00	1521.00	8050.00	7639.00	6142.00	4069.00	3858.00	3682.00	2
46	SNP_A-1715471	2721.00	280.00	298.00	351.00	3796.00	1003.00	609.00	406.00	2241.00	613.00	408.00	366.00	8145.00	1909.00	1132.00	598.00	2575.00	615.00	4
47	SNP_A-1752475	5413.00	1200.00	879.00	628.00	6187.00	2172.00	2828.00	913.00	5800.00	1957.00	1778.00	955.00	7357.00	2833.00	3538.00	1445.00	5248.00	2066.00	2
48	SNP_A-1708564	8538.00	2682.00	1315.00	628.00	10672.00	4367.00	2920.00	833.00	6963.00	1687.00	2680.00	478.00	11023.00	5461.00	2256.00	670.00	5331.00	1358.00	1

53 165 Read Ovr Block Sync Rec Caps

單一核苷酸多型性生物晶片（SNP chip）

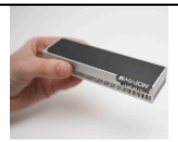


單一核苷酸與其他變異（Variant calling）




```
##fileformat=VCFv4.1
##fileDate=20090805
##source=myImputationProgramV3.1
##reference=file:///seq/references/1000GenomesPilot-NCBI36.fasta
##contig=<ID=20,length=62435964,assembly=B36,md5=f126cdf8a6e0c7f379d618ff66beb2da,species="Homo sapiens",taxonomy=x>
##phasing=partial
##INFO=<ID=NS,Number=1,Type=Integer,Description="Number of Samples With Data">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency">
##INFO=<ID=AA,Number=1,Type=String,Description="Ancestral Allele">
##INFO=<ID=DB,Number=0,Type=Flag,Description="dbSNP membership, build 129">
##INFO=<ID=H2,Number=0,Type=Flag,Description="HapMap2 membership">
##FILTER=<ID=q10,Description="Quality below 10">
##FILTER=<ID=s50,Description="Less than 50% of samples have data">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Integer,Description="Genotype Quality">
##FORMAT=<ID=DP,Number=1,Type=Integer,Description="Read Depth">
##FORMAT=<ID=HQ,Number=2,Type=Integer,Description="Haplotype Quality">
#CHROM POS     ID        REF    ALT     QUAL FILTER INFO                              FORMAT      NA00001        NA00002        NA00003
20     14370   rs6054257 G      A       29   PASS   NS=3;DP=14;AF=0.5;DB;H2           GT:GQ:DP:HQ 0|0:48:1:51,51 1|0:48:8:51,51 1/1:43:5:.,.
20     17330   .         T      A       3    q10    NS=3;DP=11;AF=0.017               GT:GQ:DP:HQ 0|0:49:3:58,50 0|1:3:5:65,3   0/0:41:3
20     1110696 rs6040355 A      G,T     67   PASS   NS=2;DP=10;AF=0.333,0.667;AA=T;DB GT:GQ:DP:HQ 1|2:21:6:23,27 2|1:2:0:18,2   2/2:35:4
20     1230237 .         T      .       47   PASS   NS=3;DP=13;AA=T                   GT:GQ:DP:HQ 0|0:54:7:56,60 0|0:48:4:51,51 0/0:61:2
20     1234567 microsat1 GTC    G,GTCT  50   PASS   NS=3;DP=9;AA=G                    GT:GQ:DP    0/1:35:4       0/2:17:2       1/1:40:3
```

全基因體定序

Platform Name	Illumina HiSeq 2500	Ion Torrent-Proton II	PacBio RS II	OxFord Nanopore Minion
Instrument				
Cost (USD) **	690 k	224 k	695 k	1 k ***
Reagent cost Per run/per GB	4126/45.84	1000/20.41	100/1111.11	900/1000
Reads per run	300 millions	280 millions	0.03 millions	0.1 millions
Average Read length	2 × 150 bp	175 bp	14,000 bp	9,000 bp
Run time	10 h	5 h	2 h	6 h
Major errors	substitution	indel	indel	deletion
Error rate (%)	0.1	1	1	4
Amplification	bridgePCR	emPCR	none, SMS	none, SMS
Advantage	low cost per GB; high output	low cost	long reads; no amplification bias	long reads; no amplification bias
Disadvantage	high cost	homopolymer errors	low throughput; high cost	high error rate

* Sources: http://www.molecularecologist.com/next-gen-fieldguide-2014/ and websites of the companies;

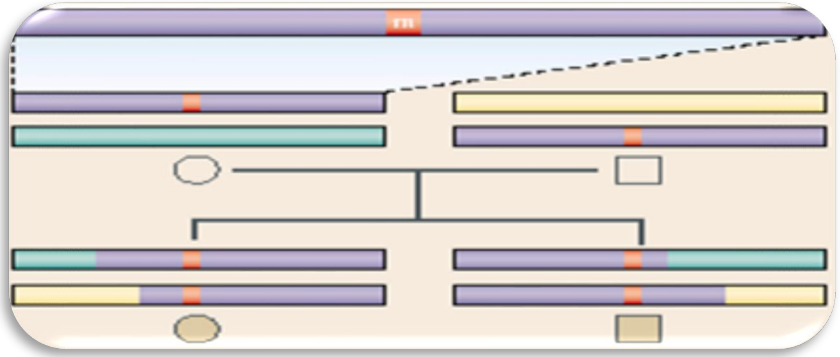
** Sources: http://www.molecularecologist.com/next-gen-table-3a-2014/;

*** Accessing fee. Sources: https://www.nanoporetech.com/products-services/minion-mki.

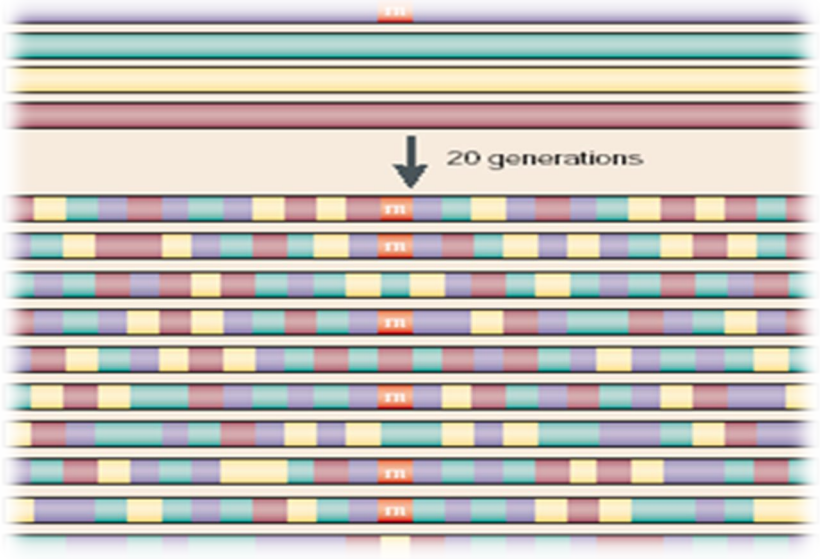
https://www.semanticscholar.org/paper/Alignment-of-Short-Reads%3A-A-Crucial-Step-for-of-in-Ye-Meehan/4963fd59b1b9295fca5f3e7568fe208679b5095b

基因研究與生物科技一日千里，窺探人類基因體的奧妙已經不再只是遙不可及的夢想。

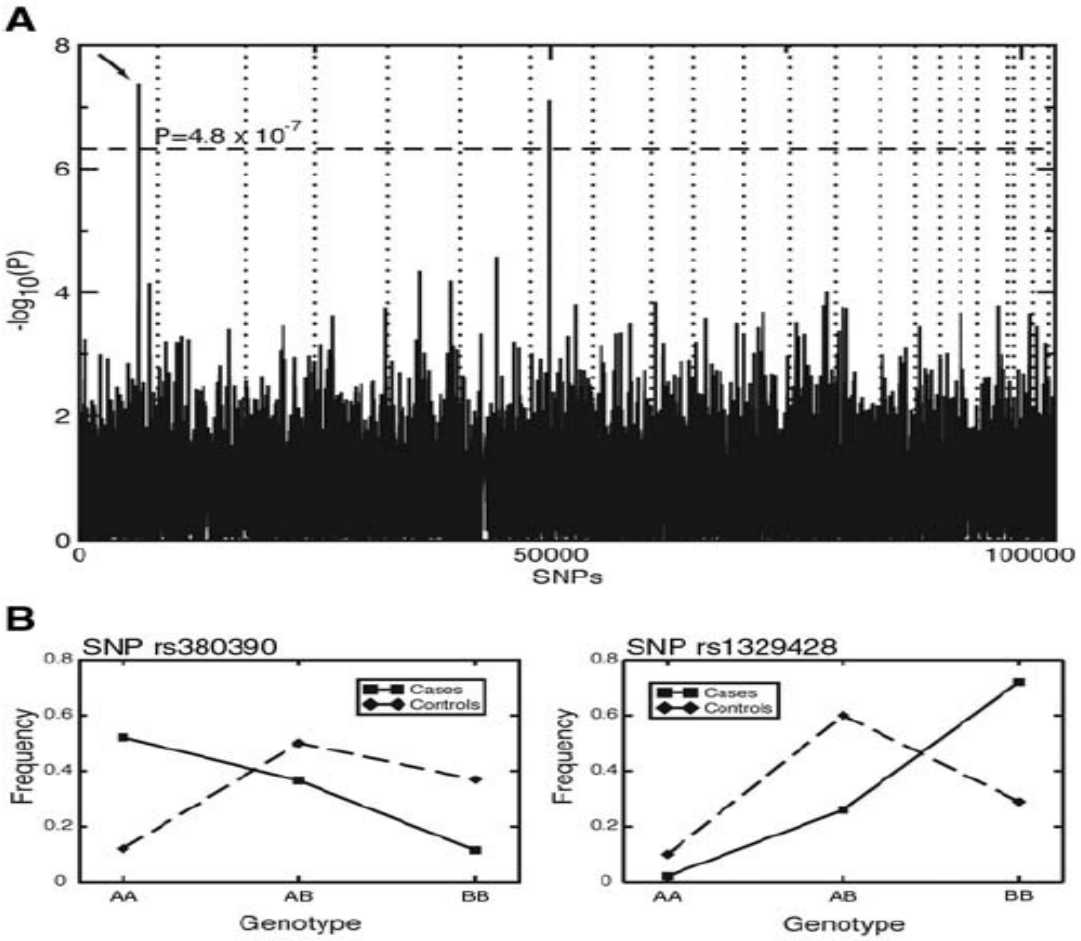
連鎖分析 (Linkage analysis)



關聯分析 (Association analysis)

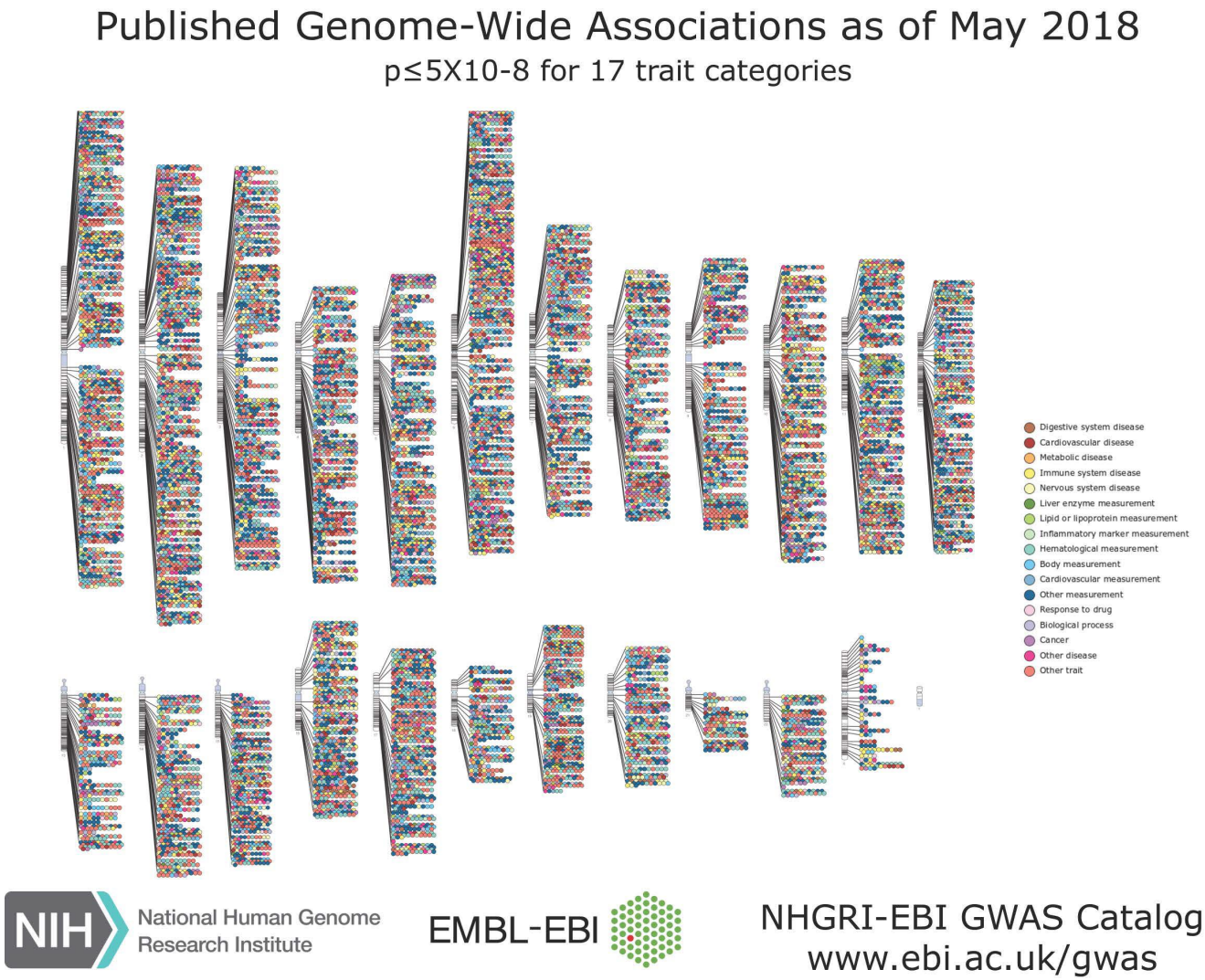


全基因體關聯性研究
(Genome-Wide Association Study, GWAS)



Klein *et al.* (2005, *Science*)

Published GWAS reports, 2005 – 2018



https://www.ebi.ac.uk/gwas/diagram

Published Genome-Wide Associations as of July 2019

p≤5X10-8 for 17 trait categories

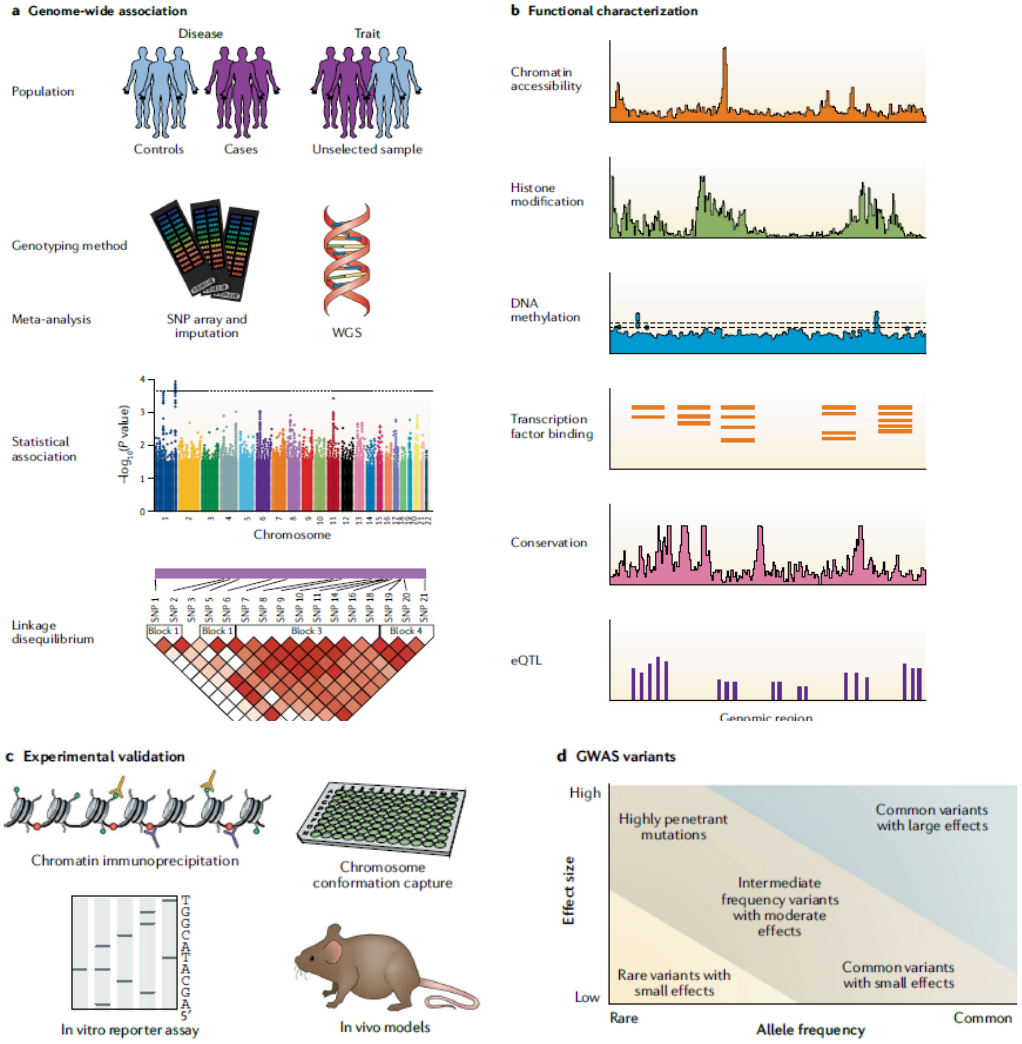
National Human Genome Research Institute

EMBL-EBI

NHGRI-EBI GWAS Catalog
www.ebi.ac.uk/gwas

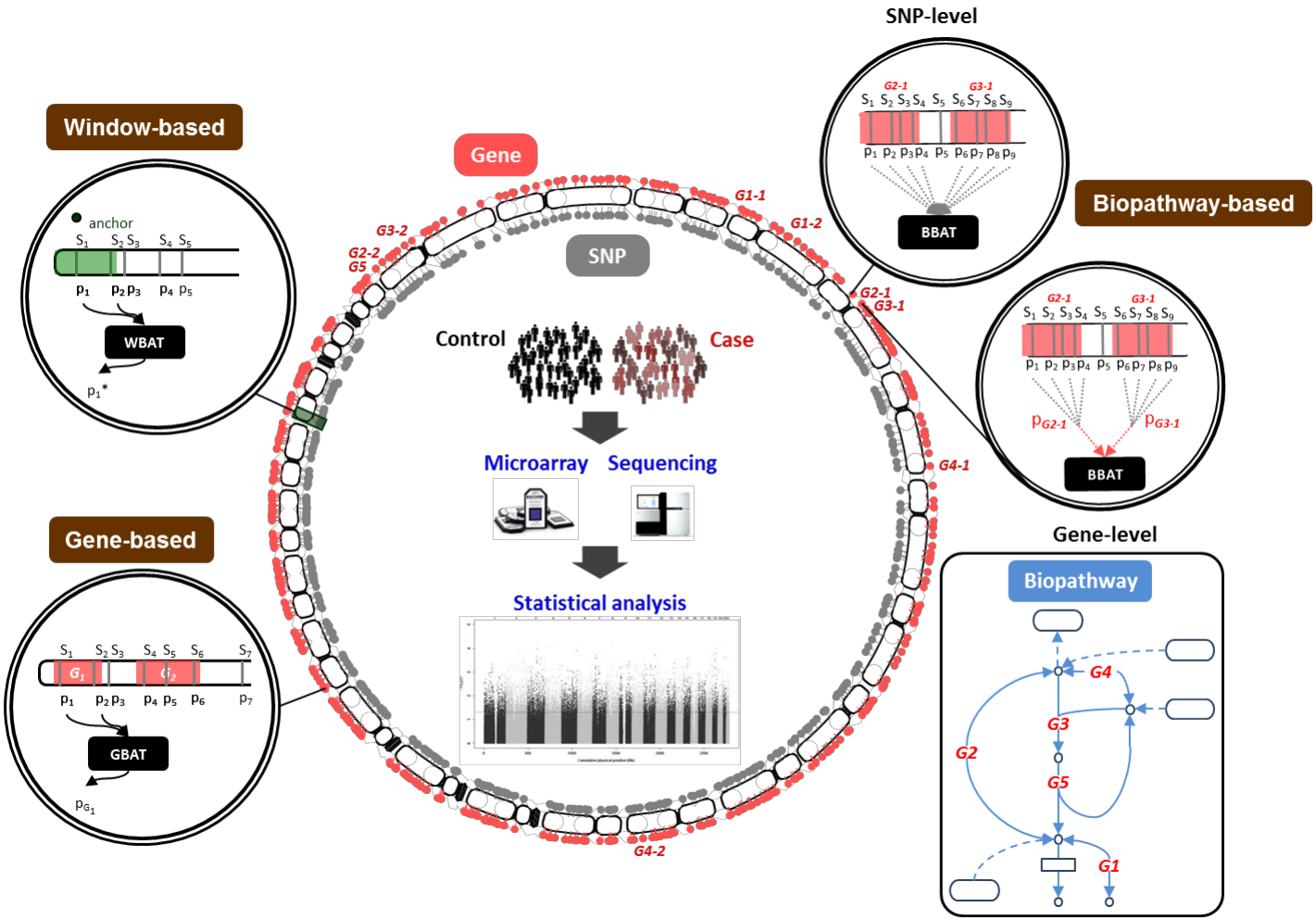
Up to Oct 2019, in 7,796 publications of GWAS, 116,109 SNPs and 159,202 SNP-disease association were identified (Genome assembly GRCh38.p12; dbSNP Build 151; Ensembl Build 96).

Genome-wide association study (GWAS)



Tam *et al*. (2019, *Nature Review Genetics*)

Genome-wide association study (GWAS)



Genome-wide phenome-wide association study (PheWAS)

Genome-wide phenome-wide association study for complex disorders & quantitative traits based on the Taiwan Biobank data

Hsin-Chou Yang[1], Chia-Wei Chen[1], Jen-Hung Wang[1], Hsiao-Chi Liao[1], Ta-Chien Chan[2], Jing-Shiang Hwang[1], Ming-Wei Su[3], Te-Chang Lee[3], Chien-Wei Lin[3], and Chun-houh Chen[1]

[1]Institute of Statistical Science, Academia Sinica, Taipei, Taiwan,

[2]Research Center for Humanities and Social Sciences, Academia Sinica, Taipei, Taiwan,

[3]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

Self-reported diseases

Category	Disease (Variable name)	Disease (Chinese name)	N (preQC)	N (postQC)
過敏	ALLERGIC_SELF	藥物過敏	2130	1869
骨科或關節疾病	OSTEOPOROSIS_SELF	骨質疏鬆症	1016	885
	ARTHRITIS_SELF	關節炎	1015	869
	GOUT_SELF	痛風	1230	1092
肺部及呼吸道疾病	ASTHMA_SELF	氣喘	768	674
	EMPHYSEMA_OR_BRONCHITIS_SELF	肺氣腫或慢性支氣管炎	331	285
心血管疾病	VALVE_HEART_DIS_SELF	瓣膜性心臟病	801	699
	CORONARY_ARTERY_DIS_SELF	冠心症	318	280
	ARRHYTHMIA_SELF	心律不整	1000	855
	CARDIOMYOPATHY_SELF	心肌症	150	128
	CONGENITAL_HEART_DIS_SELF	先天性心臟病	34	33
	OTHER_HEART_DIS_SELF	其他心臟病	39	34
	HYPERLIPIDEMIA_SELF	高血脂症	1625	1434
	HYPERTENSION_SELF	高血壓	2847	2513
	APOPLEXIA_SELF	中風	143	129
糖尿病	DIABETES_SELF	糖尿病	1125	972
消化道疾病	PEPTIC_ULCER_SELF	消化性潰瘍	3503	3069
	GERD_SELF	胃食道逆流	2867	2510
	IBS_SELF	大腸急躁症	603	530

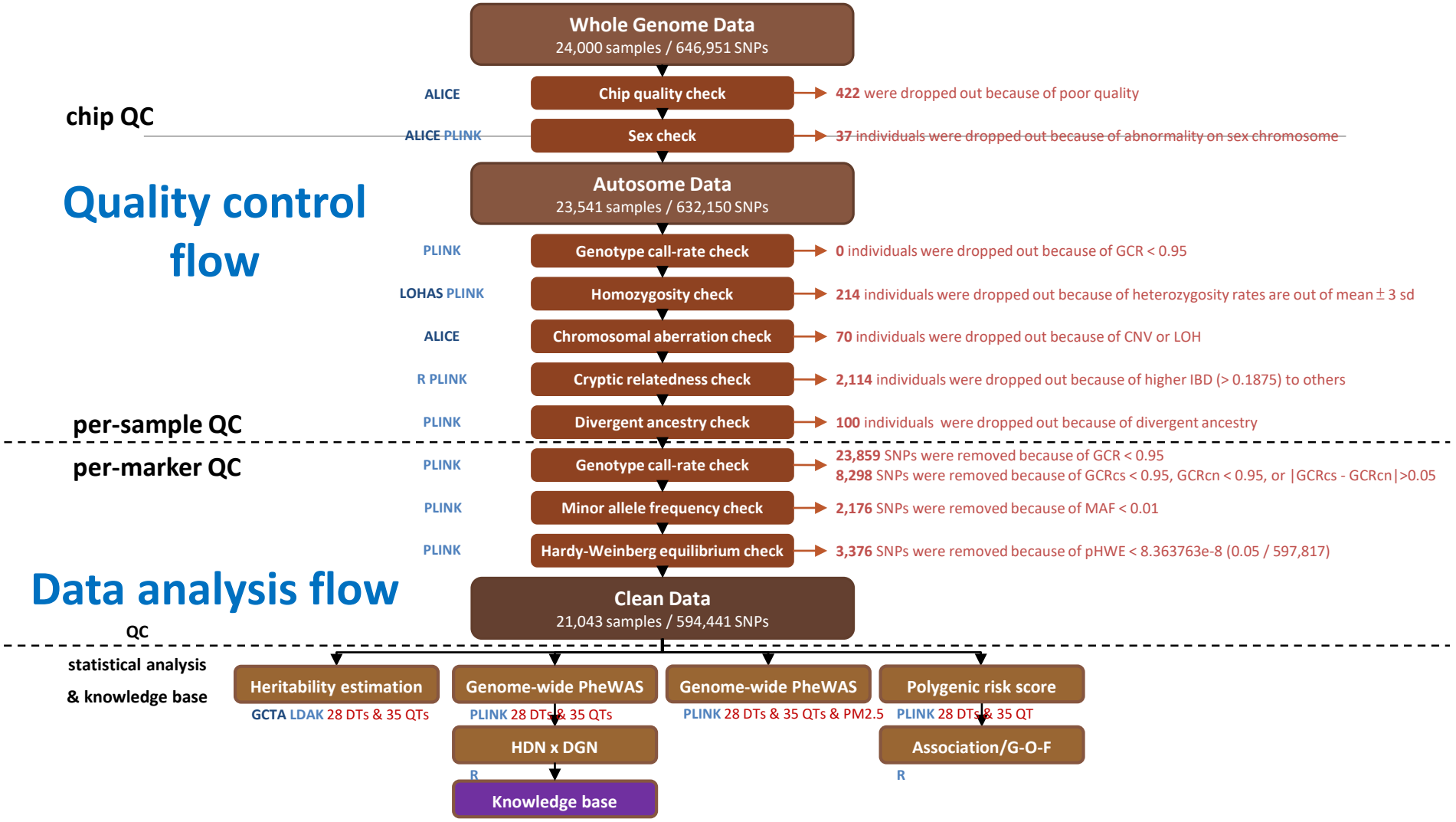
Category	Disease (Variable name)	Disease (Chinese name)	N (preQC)	N (postQC)
心理、情緒疾病	DEPRESSION_SELF	憂鬱症	774	679
	MANIC_DEPRESSION_SELF	躁鬱症	157	142
	POSTPARTUM_DEPRESSION_SELF	產後憂鬱症	63	59
	OCD_SELF	強迫症	33	28
	ALCOHOLISM_DRUG_ADDICTION_SELF	酒癮或藥物濫用	12	11
	SCHIZOPHRENIA_SELF	精神分裂症	57	44
神經系統疾病	EPILEPSY_SELF	癲癇	88	78
	HEMICRANIA_SELF	偏頭痛	593	519
	MS_SELF	多發性硬化症	5	3
	PARKISON_SELF	巴金森氏症	22	20
	DEMENTIA_SELF	失智症	6	4
其他疾病	LIVER_GALL_STONE_SELF	肝膽結石	1022	874
	KIDNEY_STONE_SELF	腎結石	1733	1533
	RENAL_FAILURE_SELF	腎衰竭	21	18
	VERTIGO_SELF	眩暈	1007	887
原位癌症	LIVER_CANCER_SELF	肝癌	0	0
	LUNG_CANCER_SELF	肺癌	0	0
	BREAST_CANCER_SELF	乳癌	0	0
	GASTRIC_CANCER_SELF	胃癌	0	0
	COLORECTAL_CAN_SELF	大腸直腸癌	0	0
	NASOPHARYNGEAL_CAN_SELF	鼻咽癌	0	0
	PROSTATE_CANCER_SELF	前列腺癌	0	0
	CANCER_OTHER_SELF	其他癌症	0	0

Quantitative traits

Category	Trait (Variable name)	Trait (Chinese name)	N (postQC)
身體檢驗	BMI	身體質量指數(BMI)	21043
	BODY_HEIGHT	身高	21043
	BODY_WEIGHT	體重	21043
	BODY_FAT_RATE	體脂肪率	20414
	BODY_WAISTLINE	腰圍	21043
	BODY_BUTTOCKS	臀圍	21043
	BODY_WB	腰臀圍比	21043
血壓與心跳	SIT_1_SYSTOLIC_PRESSURE	血壓(收縮壓)	21043
	SIT_1_DIASTOLIC_PRESSURE	血壓(舒張壓)	21043
	SIT_1_HEARTBEAT_SPEED	心跳	21043
骨質密度檢查	BONE_EXAM_RESULT	骨密度	21043
血液檢驗	RBC	紅血球 (RBC count)	21040
	WBC	白血球 (WBC count)	21040
	PLATELET	血小板 (Platelet)	21039
	HB	血紅素 (Hb)	21040
	HCT	血球比容 (HCT)	21040
血糖相關	HBA1C	醣化血色素值 (HbAlc)	21043
	FASTING_GLUCOSE	飯前血糖 (Fasting Glucose)	21043
血脂相關	T_CHO	總膽固醇 (Total cholesterol)	21043
	TG	三酸甘油脂 (Triglyceride)	21043
	HDL_C	高密度脂蛋白膽固醇 (HDL-C)	21043
	LDL_C	低密度脂蛋白膽固醇 (LDL-C)	21043

Category	Trait (Variable name)	Trait (Chinese name)	N (postQC)
肝臟相關	T_BILIRUBIN	總膽紅素 (Total bilirubin)	21043
	ALBUMIN	白蛋白 (Albumin)	21043
	SGOT	血清麩胺酸苯醋酸轉氨基酶 (SGOT)	21043
	SGPT	血清麩胺酸丙酮酸轉氨基酶 (SGPT)	21043
	GAMMA_GT	γ-麩胺醯轉移酶 (γ-Gt)	21043
肝癌指標	AFP	甲型胎兒血清蛋白 (α-fetal protein)	21043
腎臟相關	BUN	血中尿素氮 (BUN)	21043
	CREATININE	肌酸酐 (Creatinine)	21042
	URIC_ACID	尿酸 (Uric acid)	21043
	microALB	尿中微白蛋白 (Microalbumin)	21041
肝炎	ANTI_HCV_AB	C 型肝炎抗體 (HCV Ab)	21043
	HBSAG	B 型肝炎表面抗原 (HBsAg)	21043
	ANTI_HBS_AB	B 型肝炎表面抗體 (HBsAb)	21043
	ANTI_HBC_AB	B 型肝炎核心抗體 (HBcAb)	21043
	HBEAG	B 型肝炎 e 抗原 (HBeAg)	21043

Quality control flow

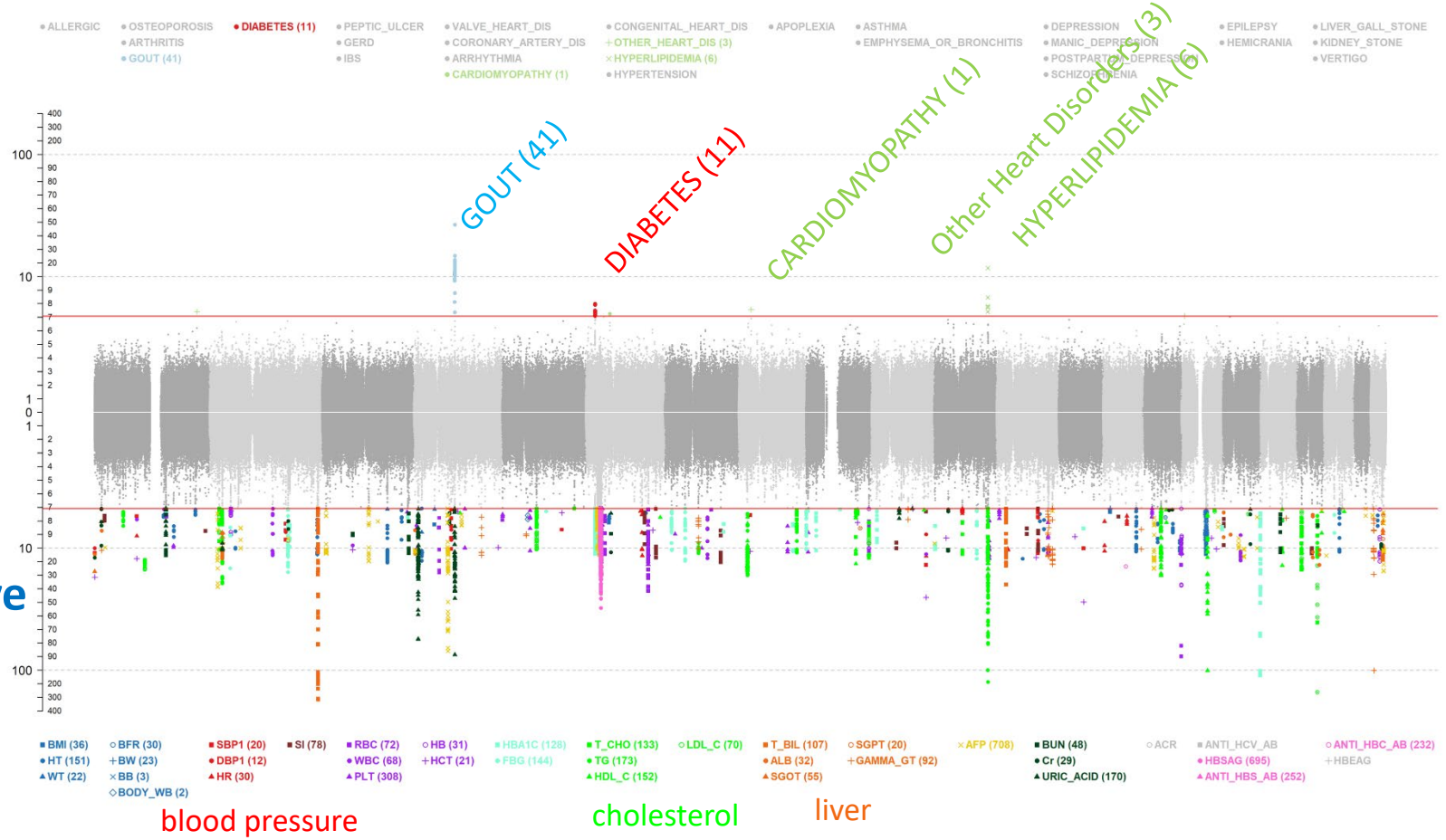


Data analysis flow

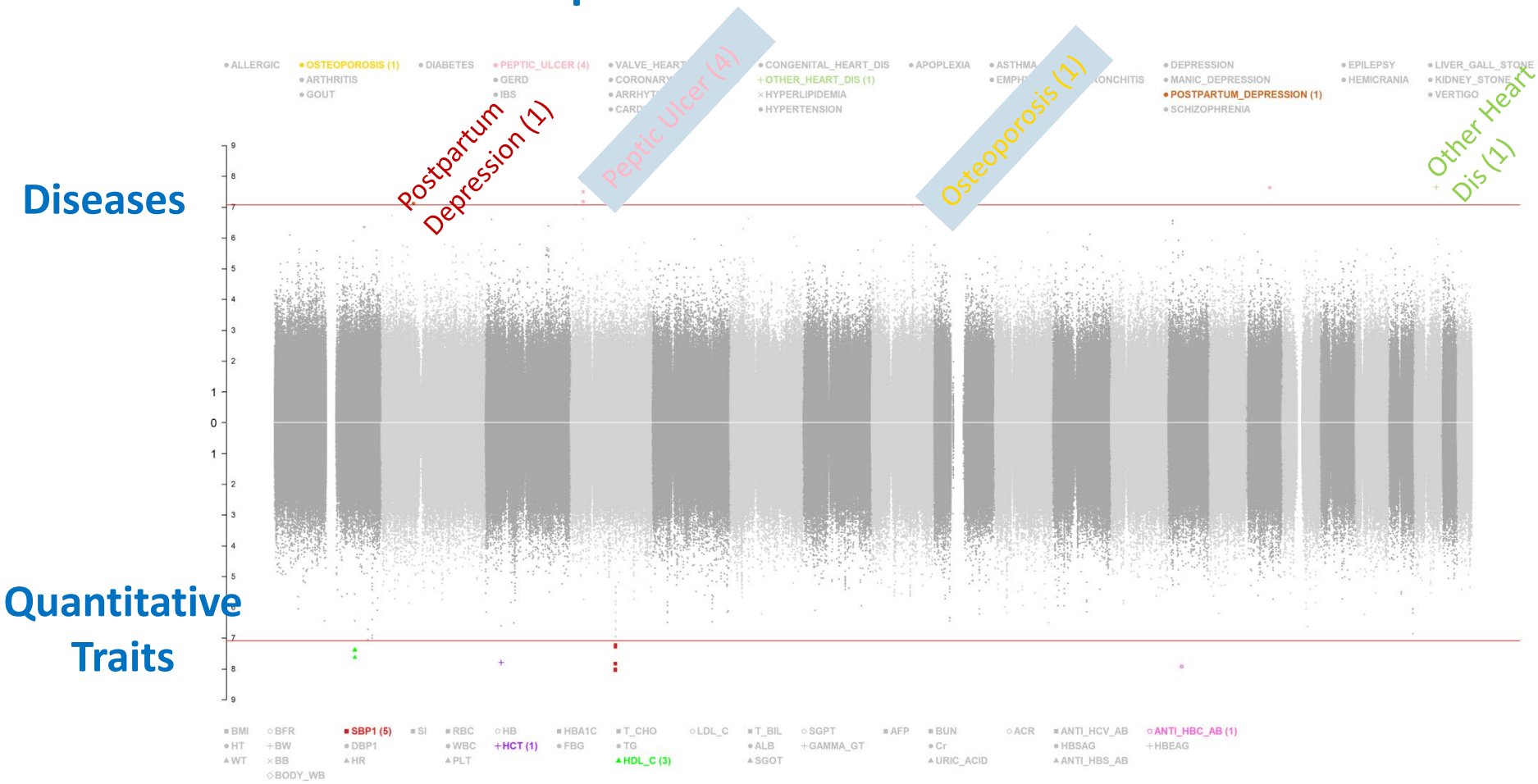
PheWAS identified important genetic variants associated with multiple diseases and quantitative traits

Self-Report Diseases

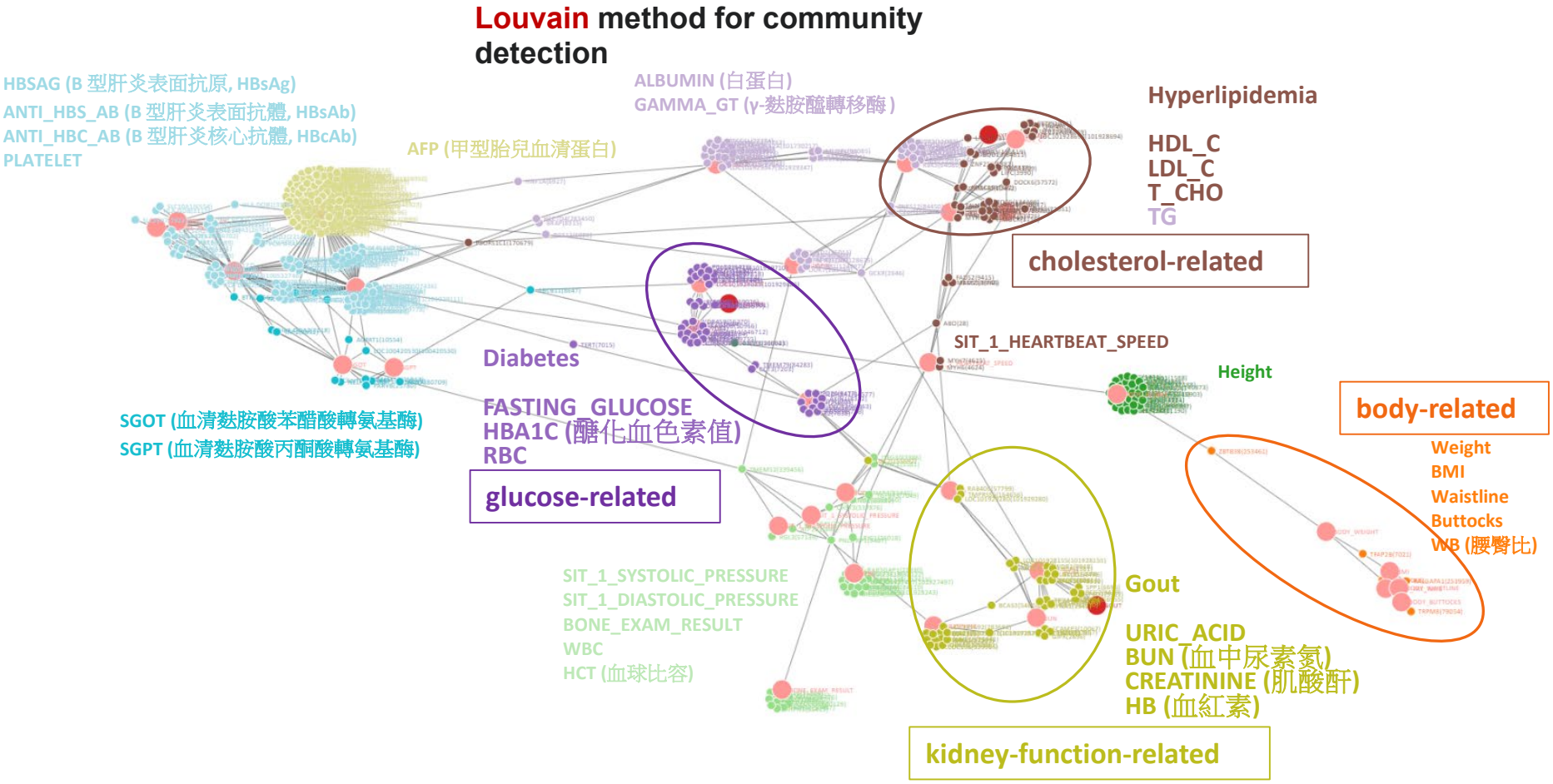
Quantitative Traits



PheWAS identified important gene-environment (PM2.5) interactions associated with multiple diseases and quantitative traits



Network of diseases and quantitative traits

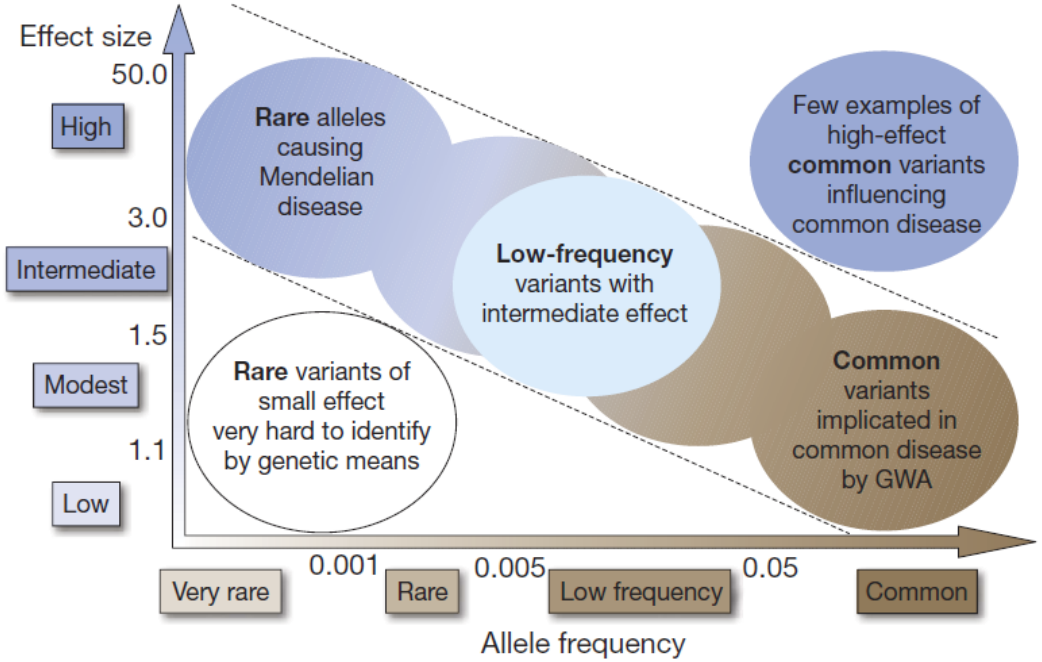


遺失遺傳度（Missing heritability）

Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration[72]	5	50%	Sibling recurrence risk
Crohn's disease[21]	32	20%	Genetic risk (liability)
Systemic lupus erythematosus[73]	6	15%	Sibling recurrence risk
Type 2 diabetes[74]	18	6%	Sibling recurrence risk
HDL cholesterol[75]	7	5.2%	Residual* phenotypic variance
Height[15]	40	5%	Phenotypic variance
Early onset myocardial infarction[76]	9	2.8%	Phenotypic variance
Fasting glucose[77]	4	1.5%	Phenotypic variance

*Residual is after adjustment for age, gender, diabetes.



Manolio *et al.* (2009, *Nature Reviews*)

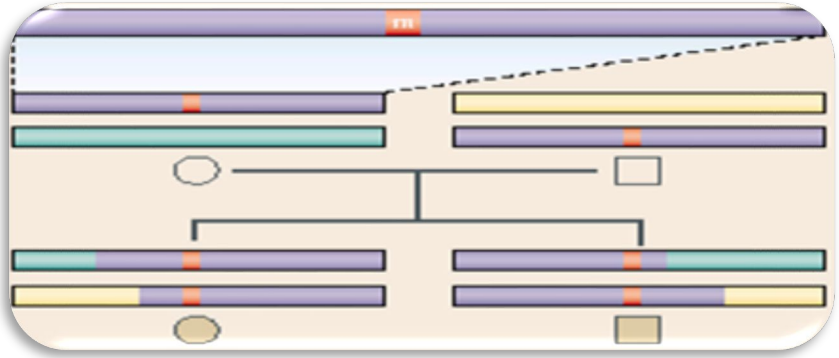
罕見變異關聯性研究
（Rare variant association analysis）

SKAT-O is the best combination of Burden Test and SKAT:

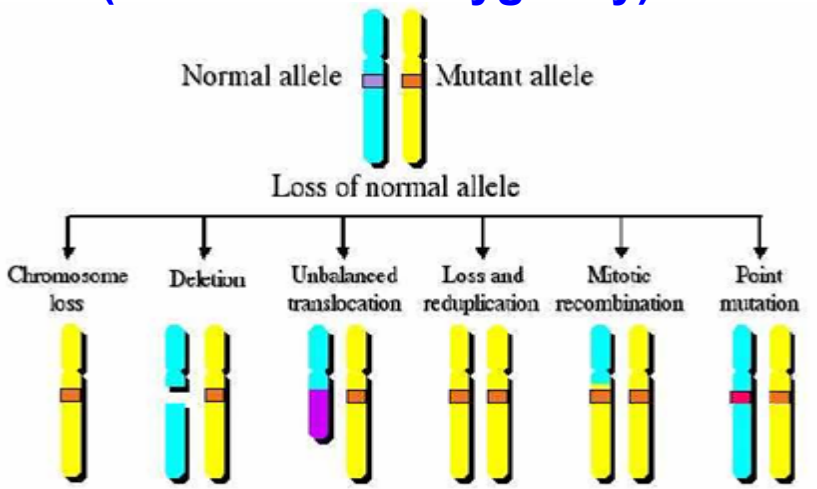
$$Q_\rho = \rho\, Q_{Burden} + (1-\rho)\, Q_{SKAT}$$

where Q_{Burden} is the statistic of the burden test and Q_{SKAT} is the statistic of the SKAT, and $0 \leq \rho \leq 1$ is coefficient that minimizes the p-value of Q_ρ.

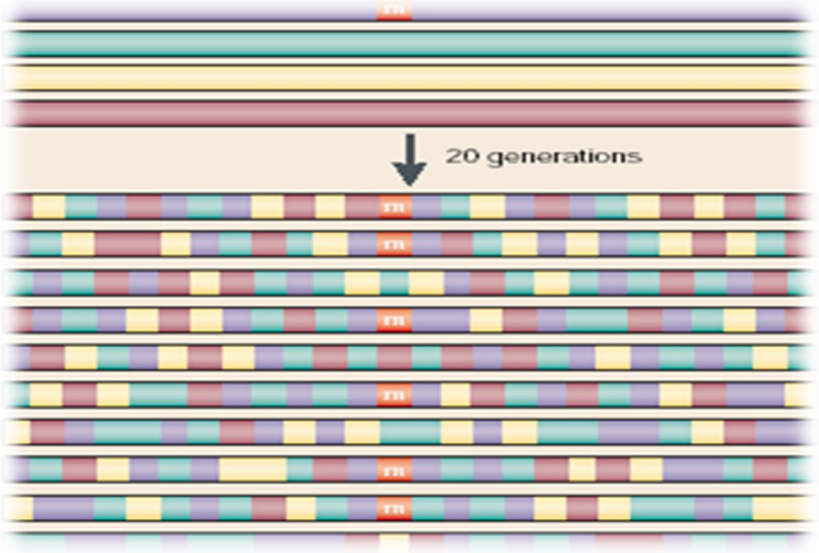
連鎖分析 (Linkage analysis)



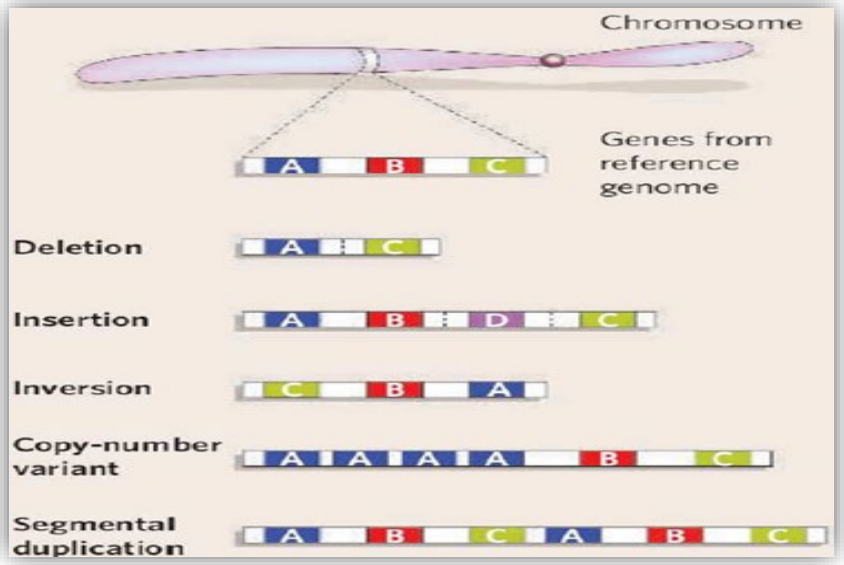
失去異合型分析 (Loss of heterozygosity)



關聯分析 (Association analysis)



拷貝數分析 (Copy number analysis)



交互作用分析
（Interaction analysis）

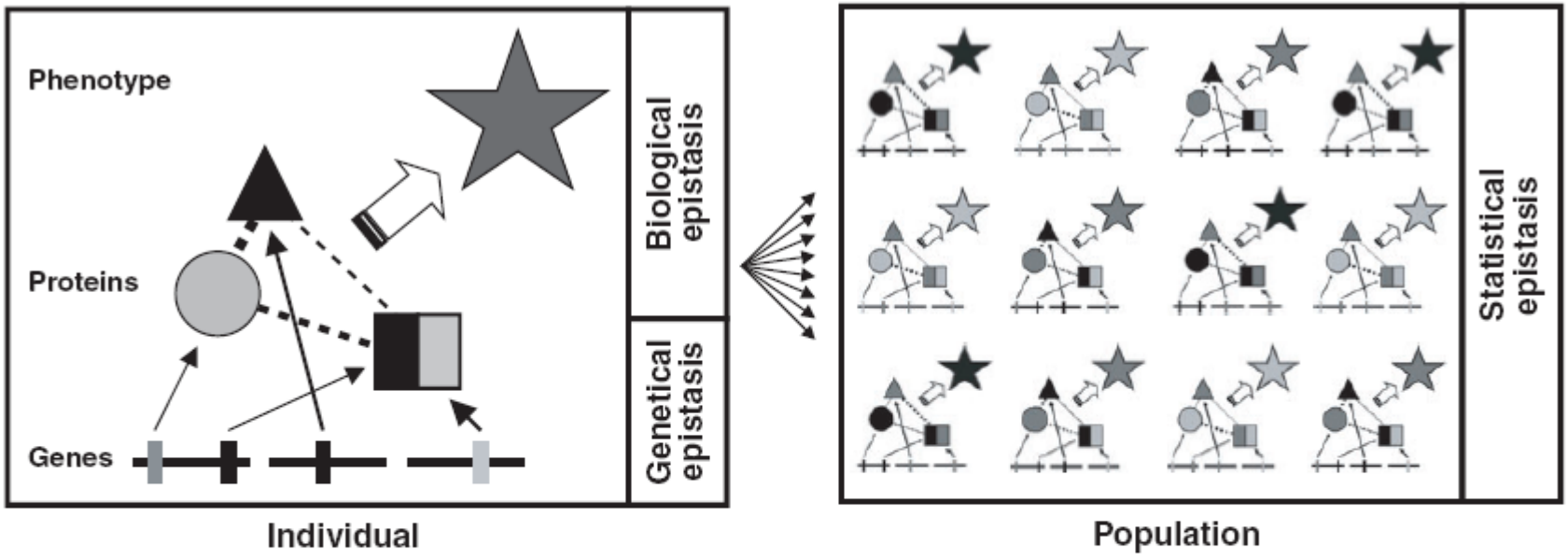
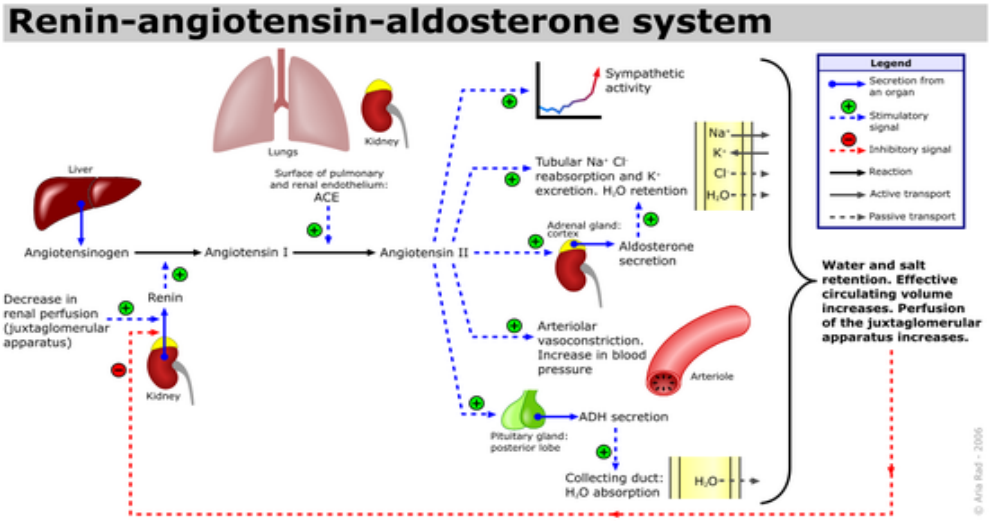
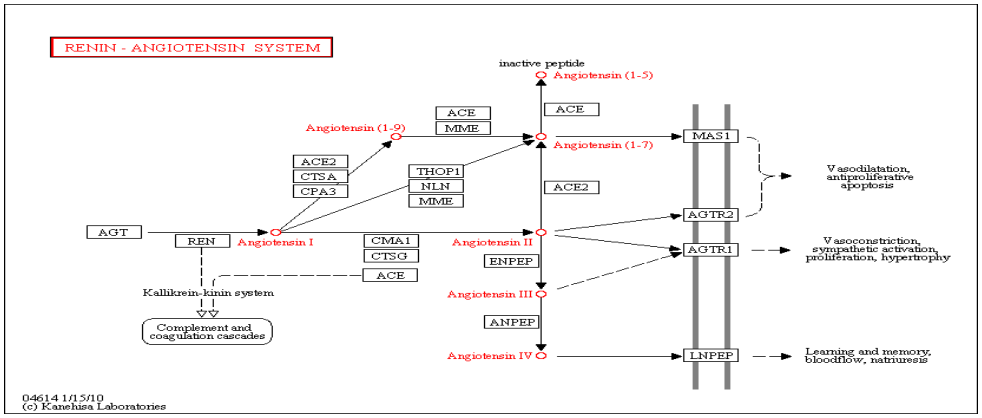


Figure 1 Genetical, biological and statistical epistasis. Genetical epistasis can be thought of as the interaction among DNA sequence variations (vertical bars) that give rise to a particular phenotype in an individual. Genetic information affects phenotype through a hierarchy of proteins (circle, square, triangle) that are involved in biological processes ranging from transcription to physiological homeostasis. The physical interactions (dashed lines) among proteins and other biomolecules and their impact on phenotype (star) constitute biological epistasis. There is a very close relationship between genetical and biological epistasis, with each occurring at the level of the individual. Differences in genetical and biological epistasis among individuals in a population give rise to statistical epistasis. It is entirely possible for genetical and biological epistasis to occur in the absence of statistical epistasis. This can happen when the DNA sequence variations and biomolecules are the same for every individual sampled from a population. Thus, genetic and biological variation is crucial for the statistical detection of epistasis. But does evidence of statistical epistasis necessarily imply genetical or biological epistasis?

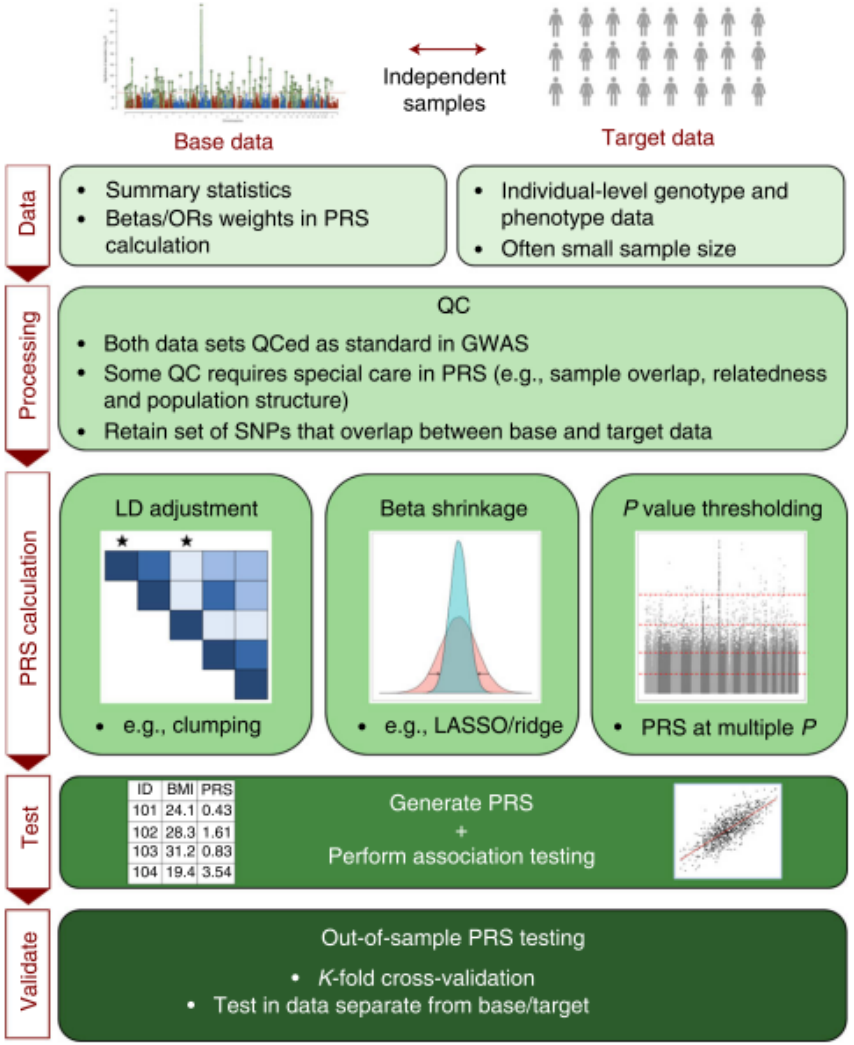
生物反應路徑分析
（Pathway analysis）



http://en.wikipedia.org/wiki/Renin-angiotensin_system

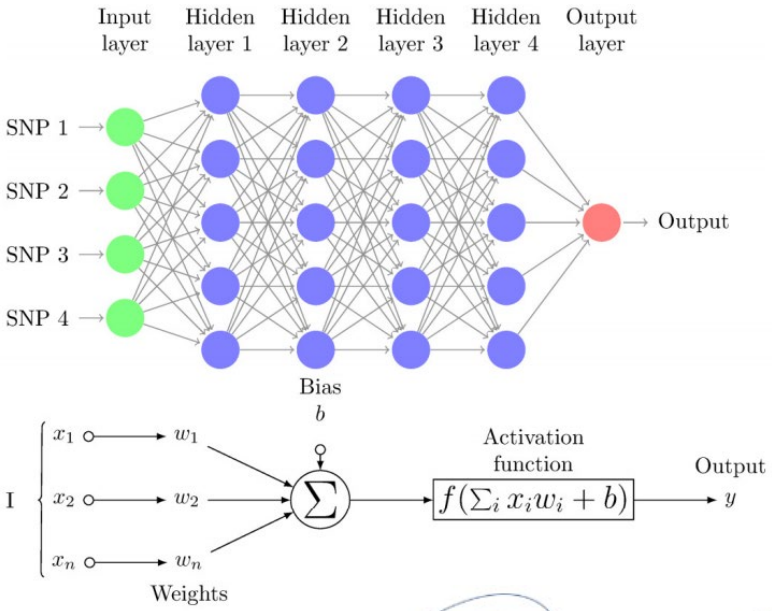


多基因風險分數
（Polygenic risk score）

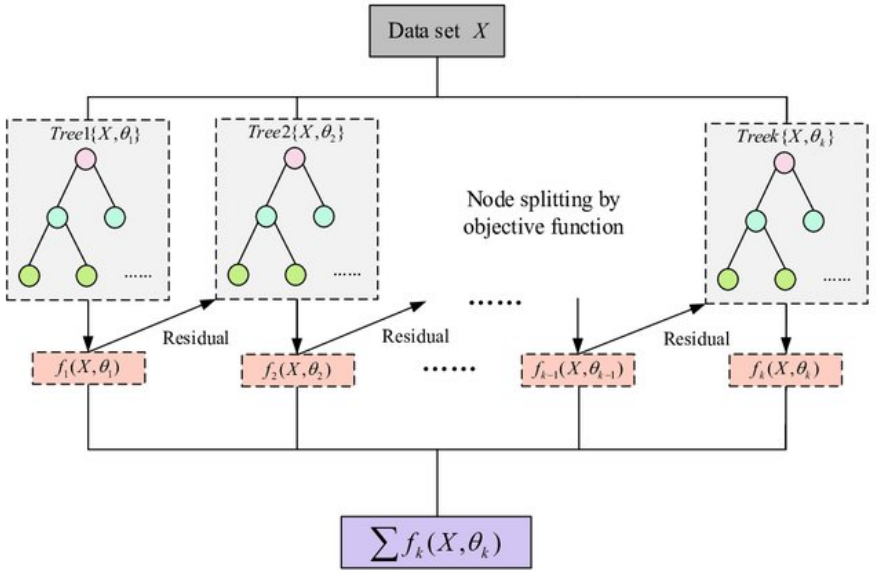


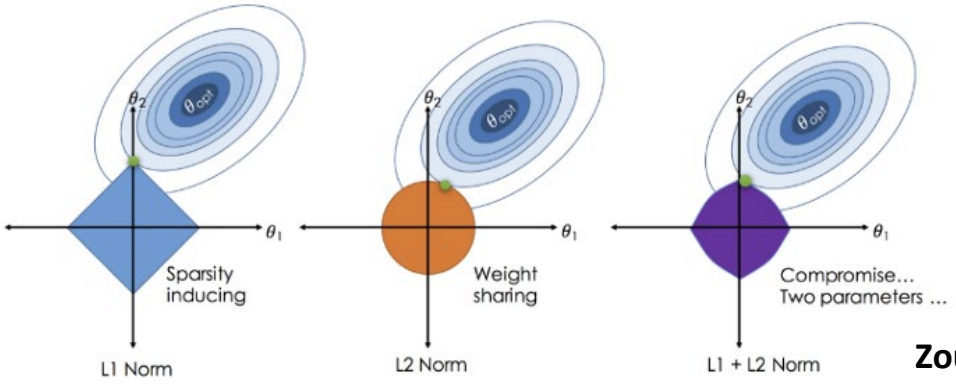
機器學習與深度學習
（Mechanism learning and deep learning）

Perez-Enciso and Zingaretti (2019, *Genes*)



Guo et al (2020, *Applied Sciences*)





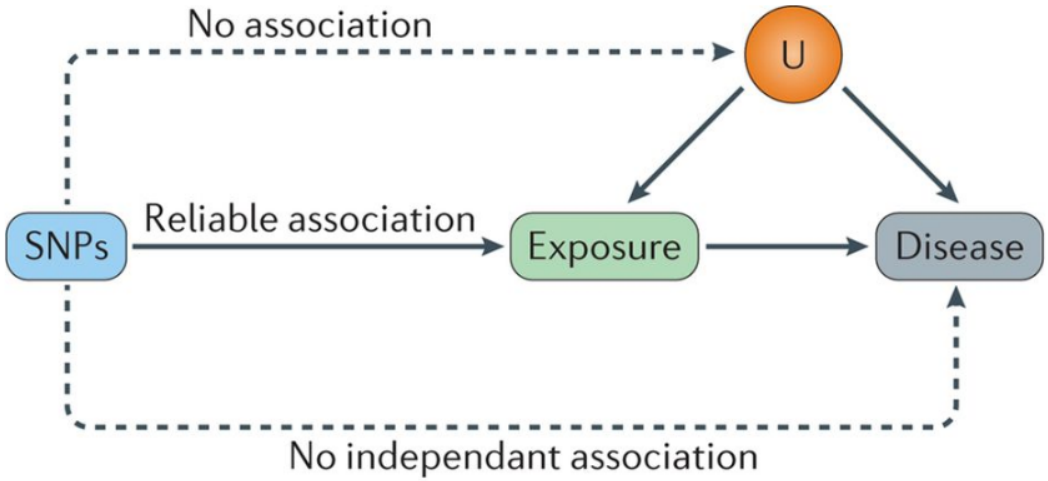
Zou and Hastie (2004, JRSS B)

孟德爾隨機化實驗
（Mendelian randomization）

MENU nature reviews cardiology

Figure 1 : Instrumental variable analysis to generate causal estimates through Mendelian randomization.

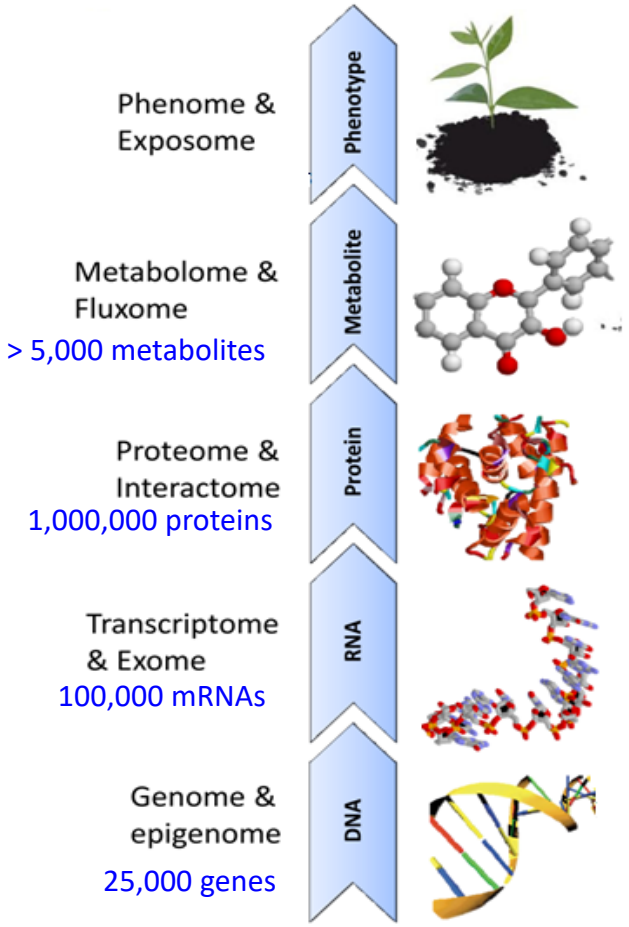
From: Mendelian randomization in cardiometabolic disease: challenges in evaluating causality



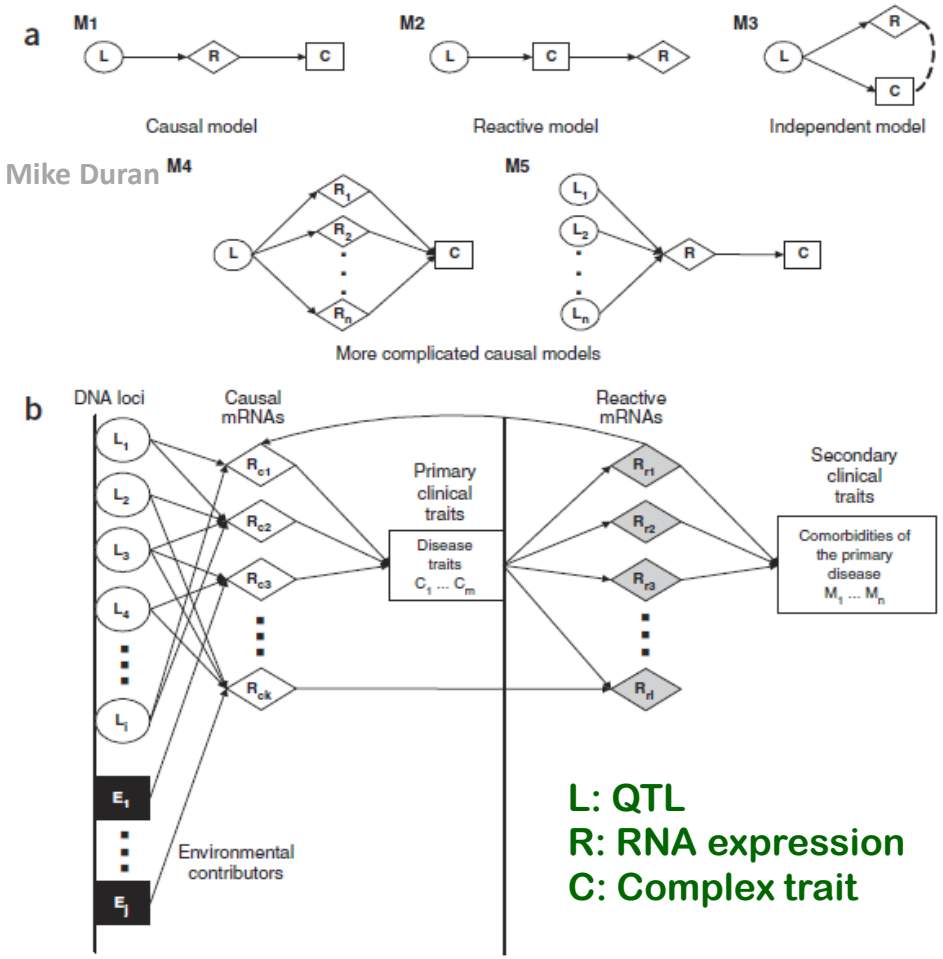
Nature Reviews | Cardiology

The three principles of instrumental variable analysis are: the instrumental variable (in this case a genetic variant either in isolation or in combination with other variants) must associate with the exposure; the instrumental variable must not associate with confounders that are either known or unknown (U); and there is no pathway from the single nucleotide polymorphism (SNP) to disease that does not include the exposure of interest. This figure is a schematic representation and should not be interpreted as a formal directed acyclic graph.

基因體系統生物分析
（Genomic convergence and -Omics）

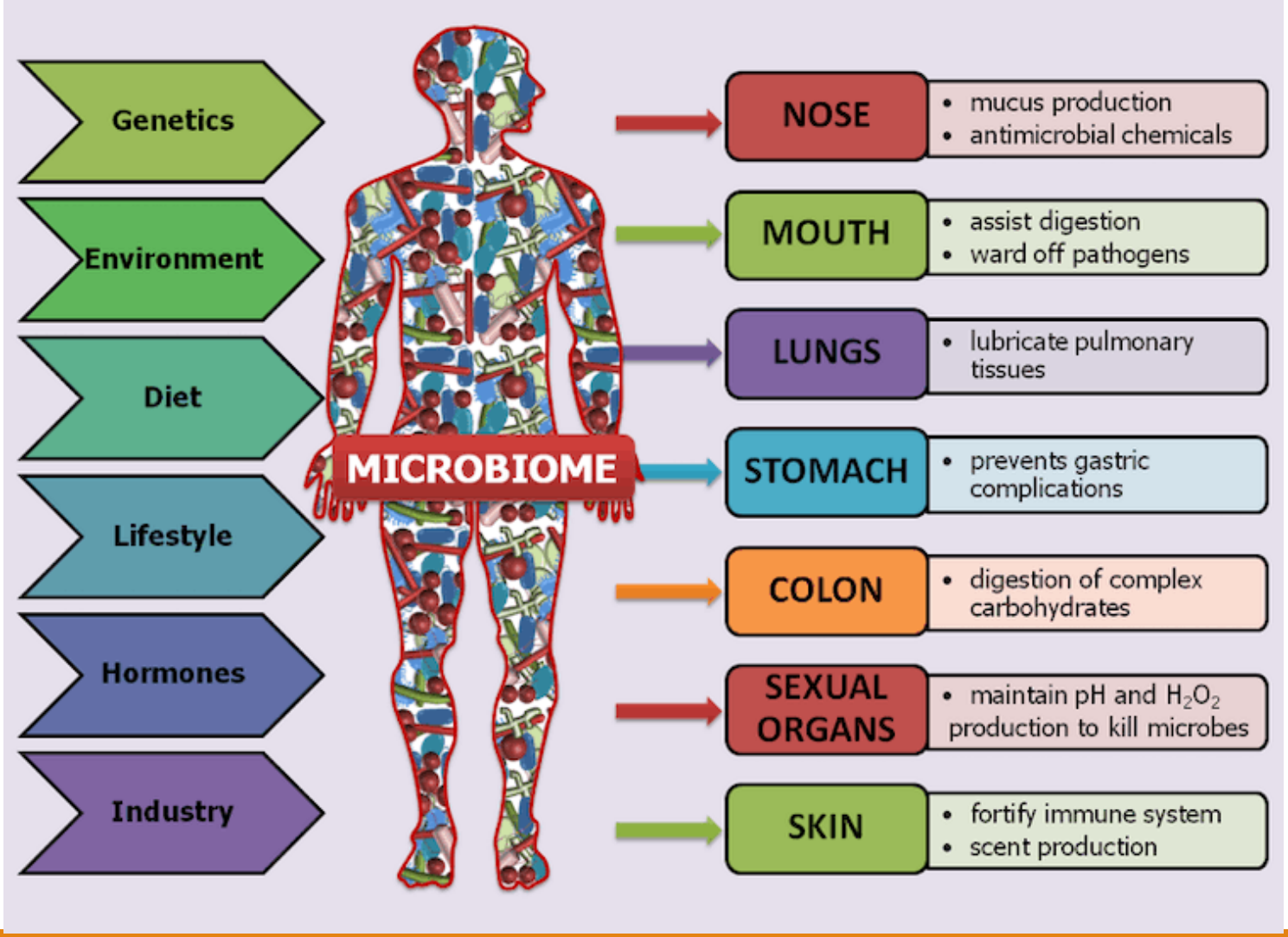


Arbona *et al.* (2013, *Int. J. Mol. Sci.*)



Schadt *et al*. (2005, *Nature Genetics*)

微生物體與總體基因體學
（Human microbial genomics and meta genomics）



https://theconversation.com/the-human-microbiome-is-a-treasure-trove-waiting-to-be-unlocked-118757

整合電子健康記錄的全表型關聯性研究（Phenome-wide association study (PheWAS) by integrating electronic health records (EHR)）

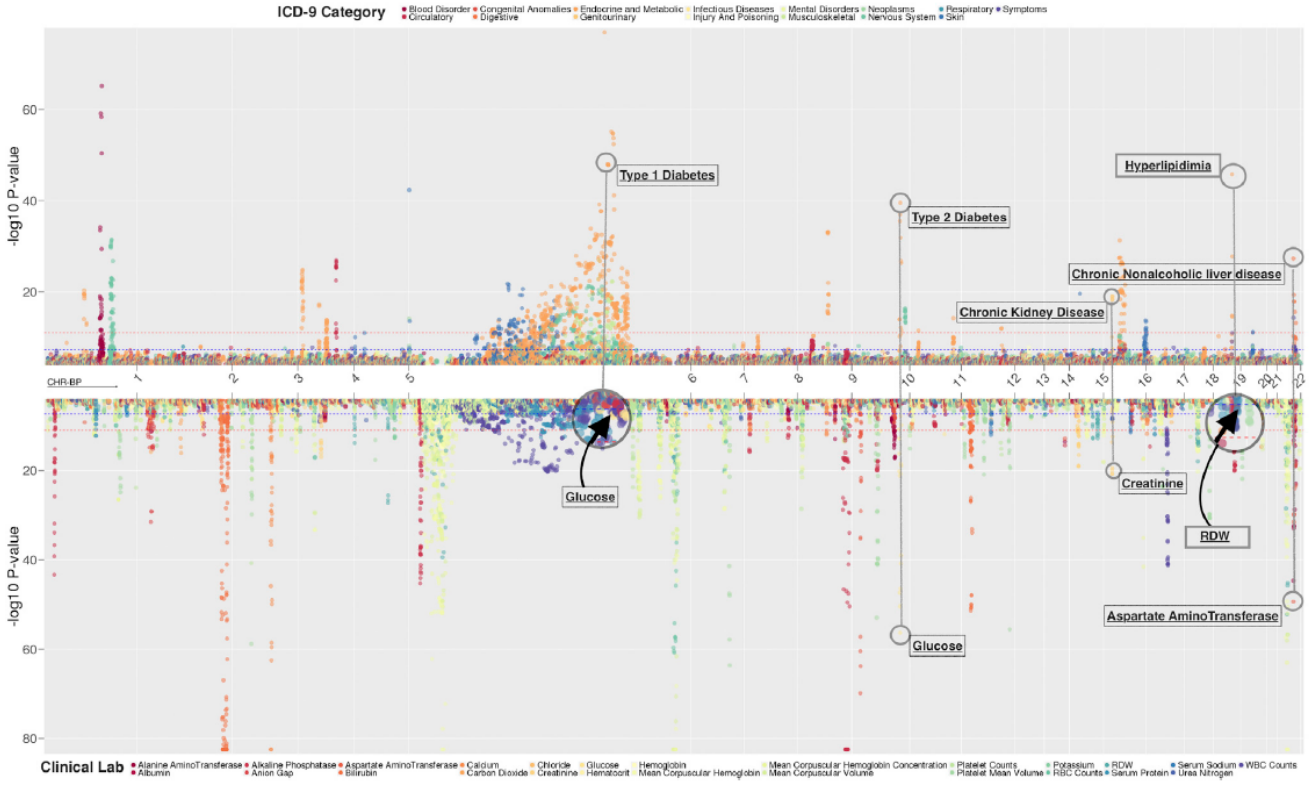


Figure 4. Integrating ICD-9 and Clinical Lab PheWAS
We present a position-by-position comparison of genetic associations the two PheWASs, one with 541 ICD-9 diagnosis codes and the other with 25 clinical laboratory measures. The horizontal axis represents genomic locations by each chromosome and the vertical axis is the $-\log_{10}$(p value) of the associations. The red and blue dotted lines are the phenome-wide significance threshold for ICD-9 and clinical lab PheWAS, respectively. We annotated examples of associations between the same SNP and highly related phenotypes across the two PheWASs.

生物資料庫研究
（Biobank studies）

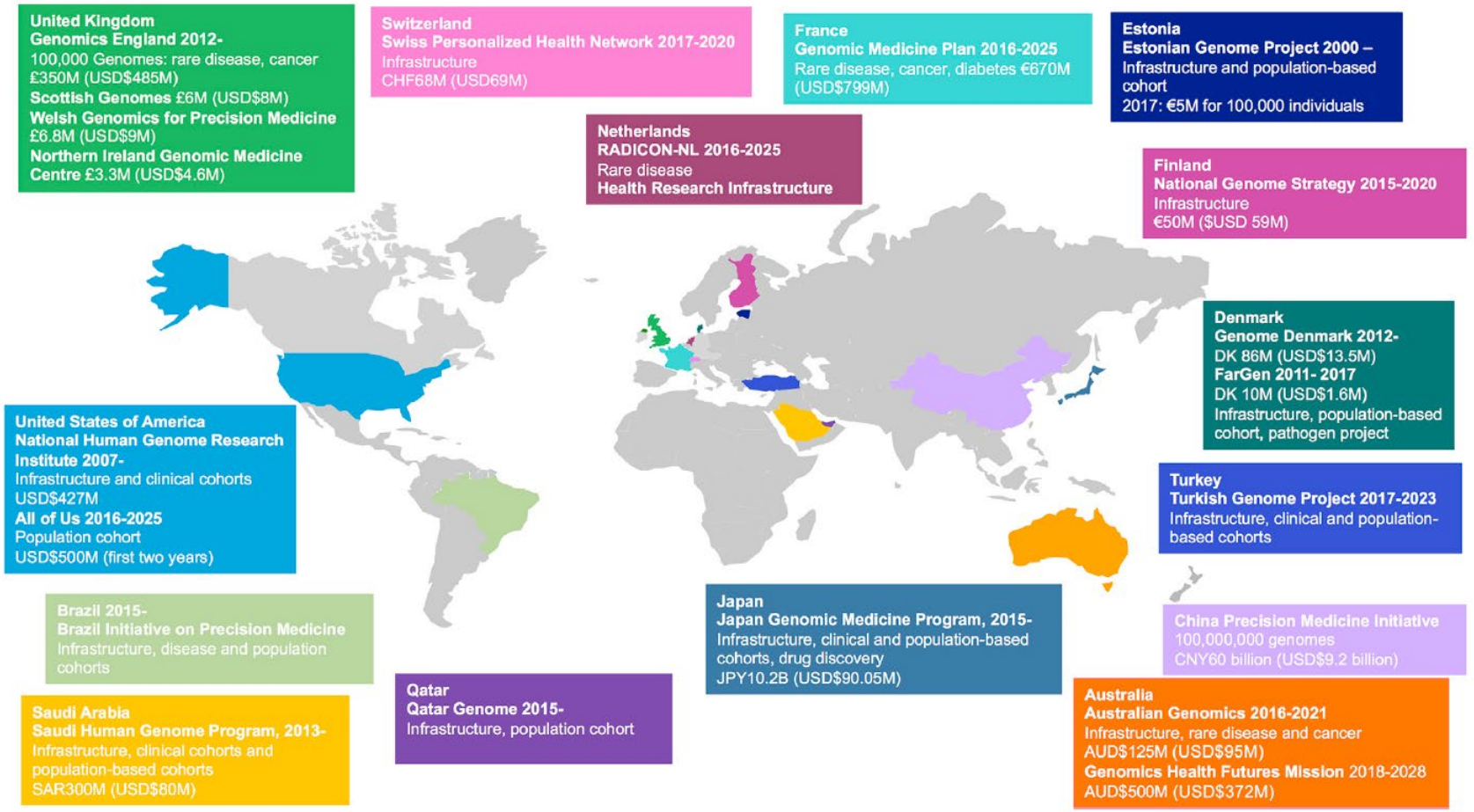


Figure 1. Map of Currently Active Government-Funded National Genomic-Medicine Initiatives

基礎統計基因體 – 線上免費課程
（Fundamental Statistical Genomics）

中文	英文	實習課	助教
人類遺傳與致病基因定位	Human genetics and disease gene mapping	R language Bioconductor	Mr. 林胤均
五個大型人類基因體研究計畫	HGP, HapMap, 1KG, TWB, and TPMI	NCBI (browser, dbSNP, PheGenI), UCSC, Taiwan View	Miss 廖筱琪
重組與連鎖分析	Recombination and linkage analysis	Pedcheck, PedigreeDraw, MERLIN, PHASE	Dr. 梁佑任
連鎖不平衡和關聯分析	Linkage disequilibrium and association analysis	HAPLOVIEW	Miss 楊智婷
全基因體關聯研究	Genome-wide association studies	PLINK	Mr. 陳佳煒
罕見變異分析	Rare variant analysis	AssotestR, SKAT, SKATO, Burden tests	Dr. 朱是鍇
基因–基因與基因–環境交互作用	Gene-gene interaction and gene-environment interaction	MDR, MDRPT, GMDR	Mr. 王仁宏

https://www.youtube.com/watch?v=IGqSaElD_T8

統計與資料科學在這些重要的遺傳研究和基因體研究工作中扮演著不可或缺的關鍵的角色

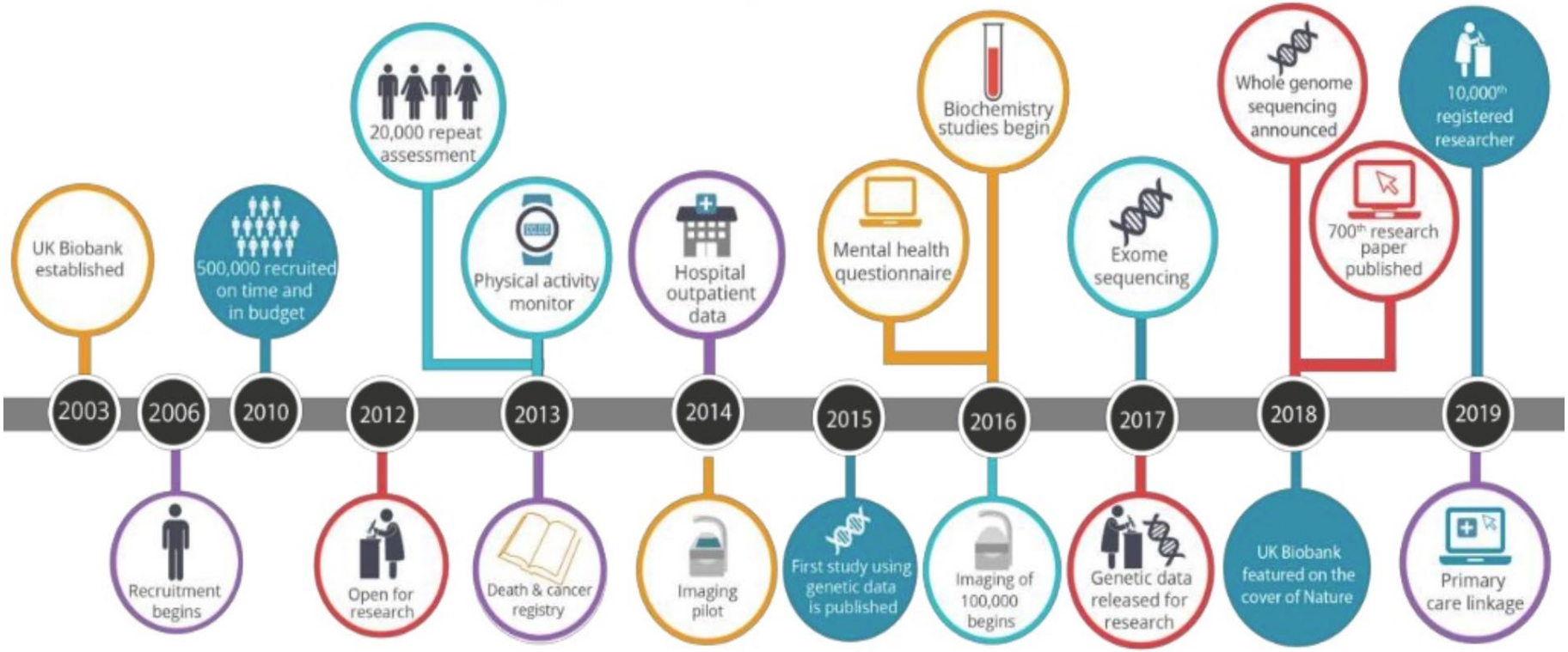
Population-based, health care system or medical center-based biobanks

TABLE 1 Description of selected major biobanks

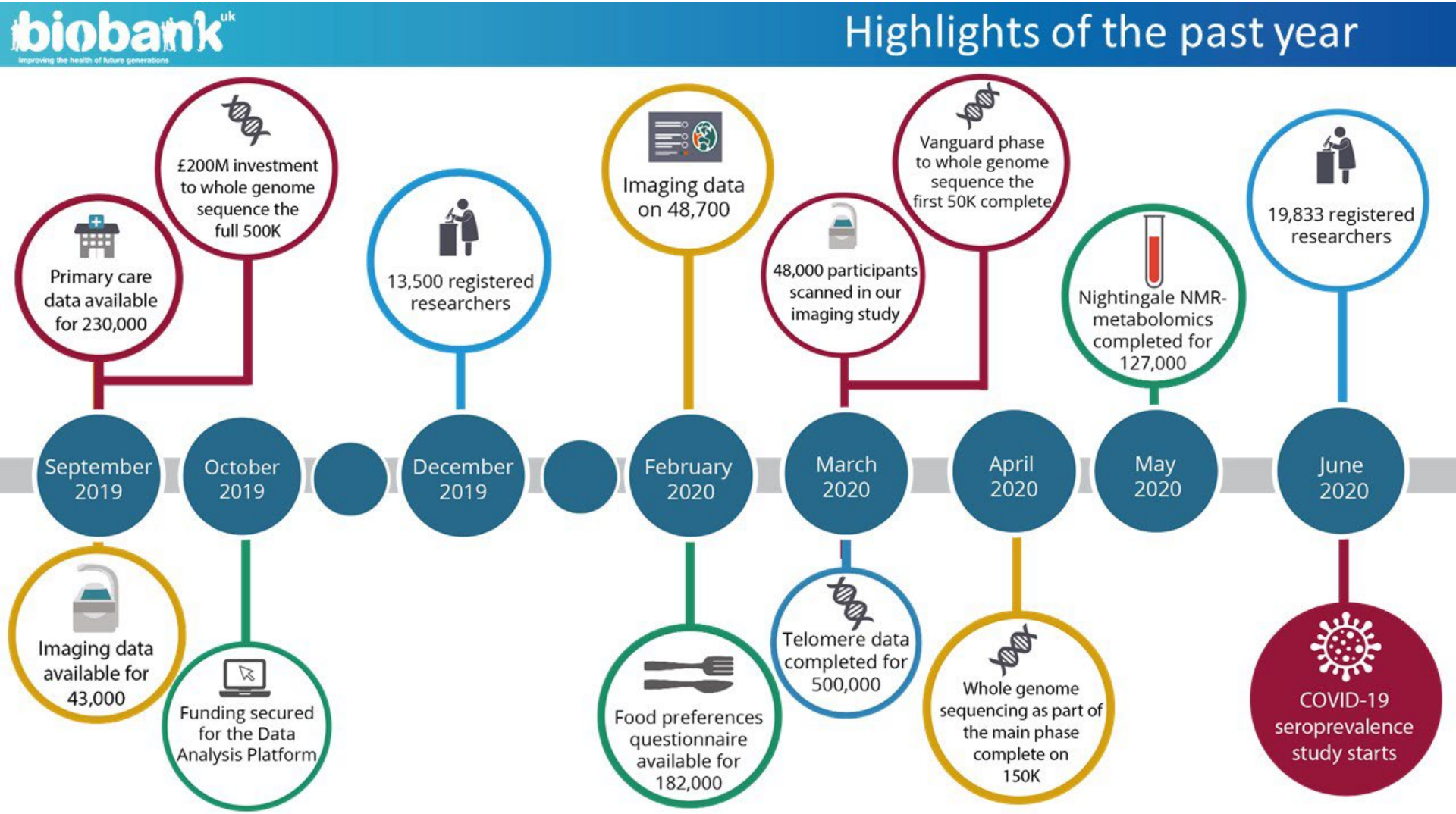
Biobank	Start year	Location	Age	Size	Type*	Institution	Access	Linked with prescriptions?	Linked to death registry?	Biospecimen collected	Survey	Website
All of Us	2018	USA	18+	1 million (goal)	Health system	National Institutes of Health	Not yet available	Yes**	—	Blood, saliva, urine	Yes	https://www.joinallofus.org/en
BioBank Japan	2003	Japan	—	200 000+	Population	Ministry of Education, Culture, Sports, Science and Technology	Inquire with biobank	—	Yes	Blood (buccal swabs or nail/hair trimmings)	Yes	http://www.ims.riken.jp/english/projects/pj02.php
BioME	—	Mount Sinai Health System	—	42 000+	Health system	Mount Sinai Health System	Inquire with biobank	—	—	Blood	Yes	https://icahn.mssm.edu/research/ipm/programs/biome-biobank
BioVU	2007	Tennessee	18+	250 000+	Health system	Vanderbilt University	Inquire with biobank	No	No	Blood	No	https://victr.vanderbilt.edu/pub/biovu/?sid=194
China Kadoorie Biobank	2004	China	30-79	510 000+	Population	University of Oxford + Chinese Academy of Medical Sciences	Application for researchers	—	Yes	Blood	Yes	http://www.ckbiobank.org/site/
deCODE Genetics	1996	Iceland	—	~500 000	Commercial	deCODE (Amgen)	Inquire with biobank	—	—	—	—	https://www.decode.com/
DiscovEHR	2014	Geisinger Health System; Regeneron Genetics Center	18+	50000	Health system	Regeneron Genetics Center + Geisinger Health System	Inquire with biobank	No	No	Blood	No	http://www.discovehrshare.com/
eMERGE Network	2007	NHGRI	All	126 000+	Network of biobanks	National Human Genome Research Institute	Application for researchers	No	No	Genetic results obtained from external sources	No	https://emerge.mc.vanderbilt.edu/
Generation Scotland	2006	Scotland	18-65	30 000+	Population	University of Edinburgh	Application for researchers	Yes	Yes	Blood, urine (saliva for some patients)	Yes	https://www.ed.ac.uk/generation-scotland

(Continues)

UK Biobank – A Prospective Cohort



Highlights of the past year



https://twitter.com/uk_biobank/status/1275434774560038915/photo/1

臺灣人體生物資料庫

宗旨：透過蒐集、儲存、系統化各式健康數據作為前哨，進而擁有高品質的健康大數據，賦予洞悉“個人化精準健康”之突破關鍵。

願景：協助專家學者解密健康密碼與改善醫療照護。

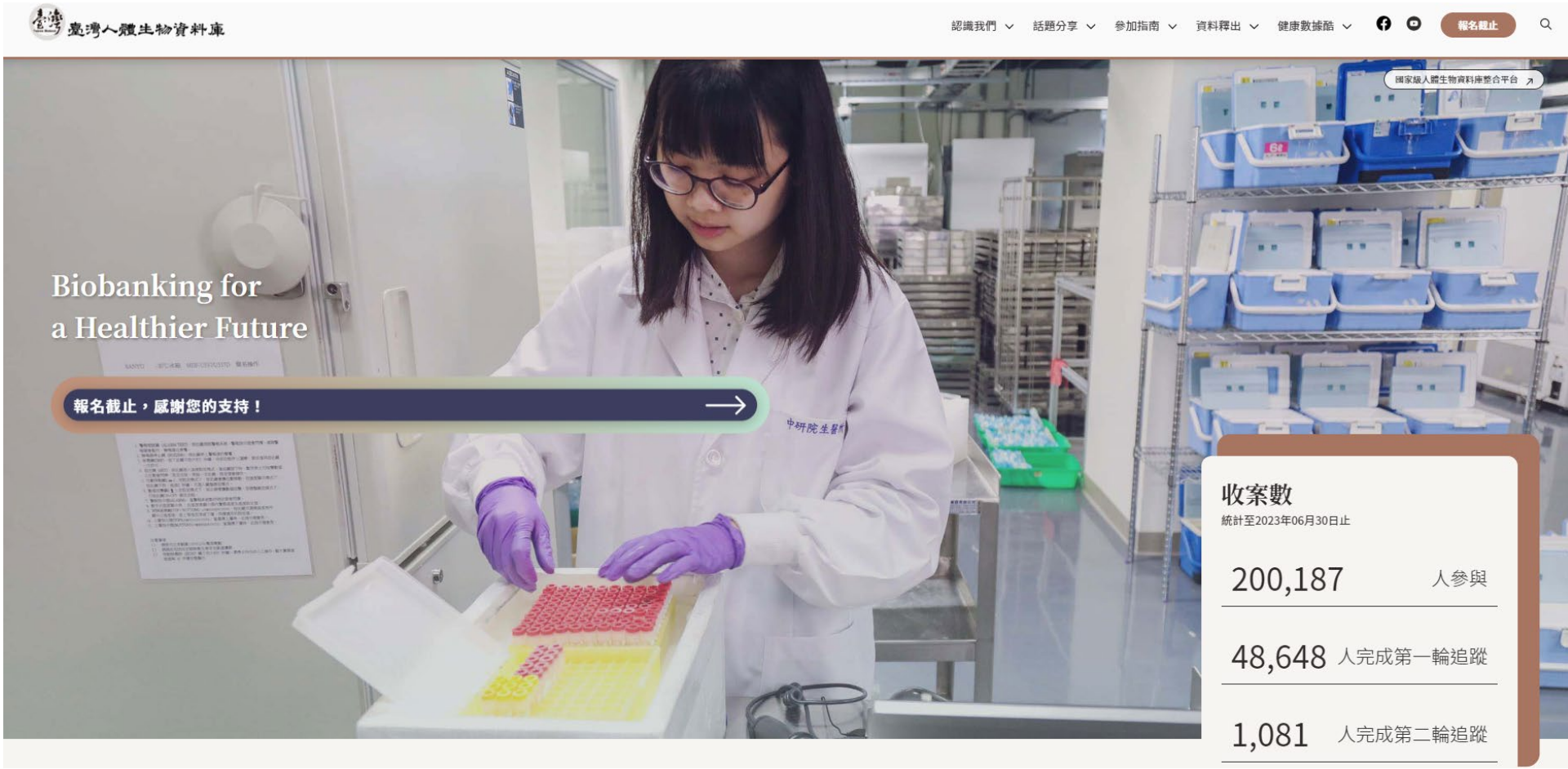
歷史：先期規劃研究計畫（2005/12～2009/10）、先期規劃之延續計畫（2010/10～2011/10）、臺灣人體生物資料庫（2012/01～至今）。

簡介：一個長期追蹤以社區族群為基礎的大型研究世代。目標收案對象為20萬名社區參與者。除收集檢體外，問卷訪查收集生活習慣、環境因子、臨床醫學資訊。另進行檢體加值，收集生物標幟資訊，建立屬於臺灣本土的人體生物資料庫，提供國內學者申請使用。

個資辨識性：去識別

研究設計：前瞻追蹤

臺灣人體生物資料庫


臺灣人體生物資料庫
認識我們
話題分享
參加指南
資料釋出
健康數據酷
報名截止
國家級人體生物資料庫整合平台
Biobanking for
a Healthier Future
報名截止，感謝您的支持！
收案數
統計至2023年06月30日止
200,187 人參與
48,648 人完成第一輪追蹤
1,081 人完成第二輪追蹤

臺灣人體生物資料庫資料型態

1. 一般參與者問卷資料
   - 1.1 ☐ A.基本人口學變項（含H.經濟狀況）
   - 1.2 ☐ B.個人健康行為
   - 1.3 ☐ C.生活環境
   - 1.4 ☐ D.飲食狀況
   - 1.5 ☐ E.家族疾病史
   - 1.6 ☐ F.女性相關問題（收案時，僅女性個案有填寫此問卷）
   - 1.7 ☐ G.MMSE（收案時，僅年齡≥60歲之個案有填寫此問卷
   - 1.8 ☐ I.中醫體質問項（收案時，由個案選擇是否填答）
2. ☐ 一般參與者身體檢測資料
   （含身高、體重、體脂肪、腰圍、臀圍、血壓、脈搏、骨密度、肺功能）
3. 一般參與者血液與尿液檢驗資料
   - 3.1 ☐ 血液學檢驗項目（含紅血球、白血球、血小板、血紅素、血球比容、醣化血色素值）
   - 3.2 ☐ 血清學檢驗項目（含飯前血糖、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）
   - 3.3 ☐ 肝膽功能類檢驗項目（含總膽紅素、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）
   - 3.4 ☐ 腎臟功能類檢驗、尿液檢驗項目（含血中尿素氮、血中肌酸酐、血中尿酸、尿中微白蛋白、尿中肌酸酐）
   - 3.5 ☐ 病毒檢驗項目（含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、D型肝炎抗體）
4. 一般參與者生物檢體
   - 4.1 DNA（以下單選，釋出溫度：4℃）
     - 4.1.1 ☐ 1 μg / 例
     - 4.1.2 ☐ 2 μg / 例
     - 4.1.3 ☐ 3 μg / 例
   - 4.2 血漿（以下單選，釋出溫度：-80℃）
     - 4.2.1 ☐ 1 管 / 例
     - 4.2.2 ☐ 2 管 / 例
   - 4.3 尿液（以下單選，釋出溫度：-80℃）
     - 4.3.1 ☐ 1 管 / 例
     - 4.3.2 ☐ 2 管 / 例
5. 實驗資訊
   - 5.1 全基因體定型資料
     - 5.1.1 ☐ TWBv1.0 晶片
     - 5.1.2 ☐ TWBv2.0 晶片
     - 5.1.3 ☐ TWBv1.0 晶片或 TWBv2.0 晶片
   - 5.2 全基因體定序資料
     - 5.2.1 ☐ Illumina平台
     - 5.2.2 ☐ Ion Proton平台
     - 5.2.3 ☐ Illumina 平台或 Ion Proton 平台
   - 5.3 ☐ 全基因體甲基化晶片資料
   - 5.4 ☐ 人類白血球組織抗原分型資料
   - 5.5 ☐ 血液代謝體資料
   - 5.6 尿液環境物質代謝物含量資料
     - 5.6.1 ☐ 尿液塑化劑含量資料
     - 5.6.2 ☐ 尿液三聚氰胺含量資料

臺灣人體生物資料庫檢體加值與收費

實驗加值人數

全基因體定型資料（TWBv1.0）	27719
全基因體定型資料（TWBv2.0）	88343
全基因體定序資料（WGS）	2010
基因體甲基化晶片資料（MET）	2474
人類白血球組織抗原分型資料（HLA）	1102
血液代謝體資料（MTBB）	869
塑化劑（PAEs）	1353
三聚氰胺（Melamine）	1353
影像資料集（MI）	0

檢體庫存量

DNA(μg)	2415296
Plasma(管，每管 0.4 ml)	892003
Urine(管，每管 1.0 ml)	510004

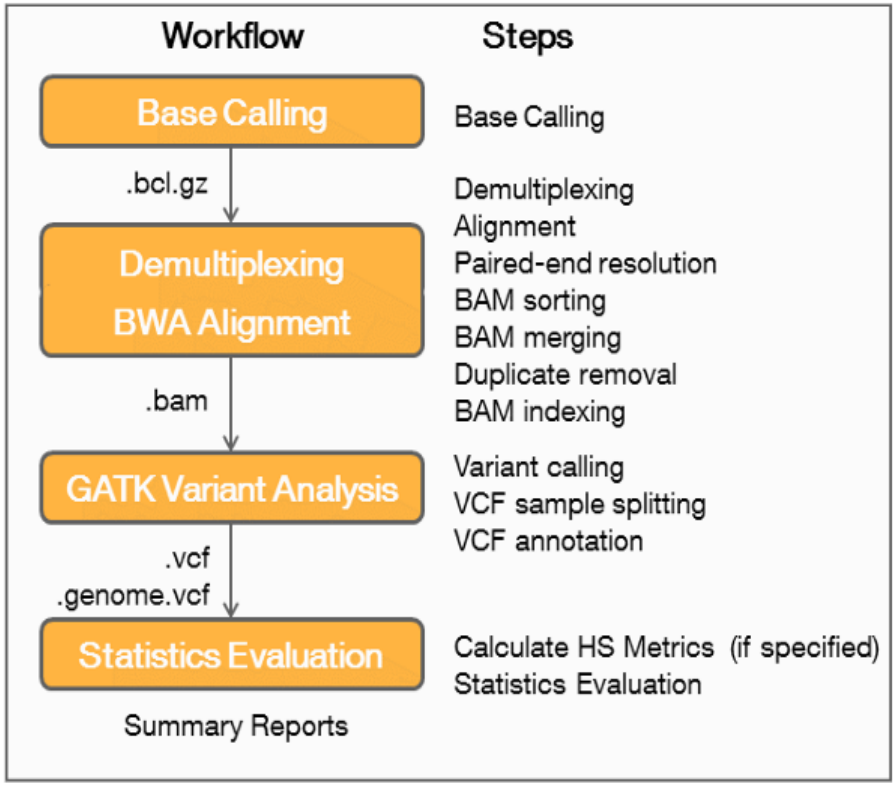
二、以整批數位資料集申請者：

一般參與者問卷、身體檢測及血液尿液檢驗數位資料集 166,457筆	一般參與者全基因體定型數位資料集 114,600筆	一般參與者全基因體定序數位資料集 2,010筆
NT$50,000 每數位資料集	NT$20,000 每數位資料集	NT$5,000 每數位資料集
一般參與者全基因體甲基化晶片數位資料集 2,474筆	一般參與者人類白血球組織抗原分型數位資料集 1,102筆	一般參與者進階追蹤數位資料集(影像資料) 20,408筆
NT$3,000 每數位資料集	NT$3,000 每數位資料集	NT$30,000 每數位資料集

臺灣人體生物資料庫檔案大小

基因型鑑定資料 **GWAS**： CEL 檔: 30MB/人
全基因型定序資料 **WAS**： Thermo Fisher: BAM file: 400G/人 VCF file: 500MB/人 全基因體定序資料 **WGS**： Illumine: Fastq file: 80G/人 BAM file: iSAAC pipeline 50G/人、GATK pipeline 150G/人 VCF file: iSAAC pipeline 5G/人、GATK pipeline 5G/人
甲基化晶片資料 **MET**： txt and idat file: 100MB/人
人類白血球組織抗原分型資料 **HLA**： csv and BAM file: 400MB/人
血液代謝體資料 **MTBB**： FID and SER file: 15MB/人

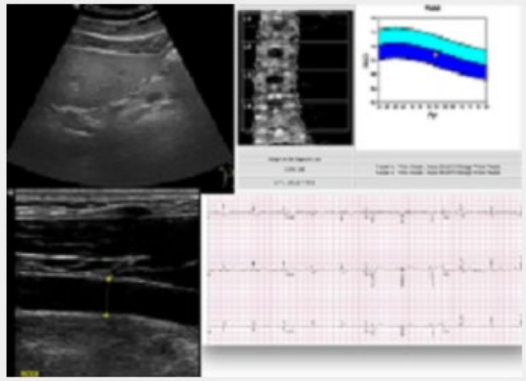
Figure 1 Enrichment Workflow Diagram



臺灣人體生物資料庫進階追蹤資料

進階追蹤醫學影像資料
Medical Images in 1st VIP Follow-up Stage

臺灣人體生物資料庫2016年正式開始「進階追蹤」收案，即除了一般追蹤的收案項目（問卷、身體檢測、血液尿液檢驗，以及抽血留尿等）以外，外加委託淡水馬偕醫院影像車至收案駐站進行腹部超音波、全身骨密度、心電圖、頸動脈超音波等醫學影像檢查。除了釋出DICOM原檔之外，亦附上無需安裝之簡易閱讀軟體CDView與使用說明，以及儀器紀錄之數值或判讀結果表格，提供給申請人使用。



實驗加值人數

全基因體定型資料（TWBv1.0）	0
全基因體定型資料（TWBv2.0）	0
全基因體定序資料（WGS）	0
基因體甲基化晶片資料（MET）	0
人類白血球組織抗原分型資料（HLA）	0
血液代謝體資料（MTBB）	0
塑化劑（PAEs）	0
三聚氰胺（Melamine）	0
影像資料集（MI）	22071

檢體庫存量

DNA(μg)	713025
Plasma(管，每管 0.4 ml)	252362
Urine(管，每管 1.0 ml)	142790

影像資料集（MI）

- 全身骨密度
  (Bone Mass Density, BMD)
- 心電圖
  (Electrocardiography, EKG)
- 腹部超音波
  (Abdominal Ultrasonography, ABD)
- 頸部超音波
  (Vertebral Artery Ultrasound, VAS)
- 甲狀腺超音波
  (Thyroid Gland Ultrasound, THY)

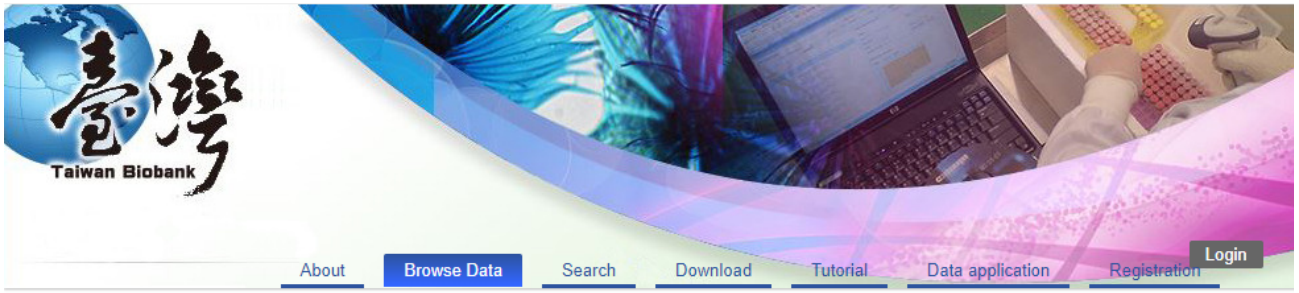
https://www.biobank.org.tw/about_value.php

TWB1.0 和 TWB2.0 生物晶片

TWB 2.0Array Plate的比較如下：

	TWB	TWB 2.0
Price	學術界：NT$350,075/plate 產業界：NT$427,785/plate (95 samples /plate)	學術界：NT$96,000/plate 產業界：NT$96,000/plate (96 samples/plate)
SNP probes	653,291	752,921
SNPs挑選之來源	● CHB Array，共525,652個SNP點。 ● 已發表之多項癌症相關GWAS研究中，具有統計顯著意義之SNP點。 ● 本中心過去使用多種類型晶片所得的結果，挑選在國人樣本中具有多型性（polymorphism）之SNP點。 ● 採用whole exome sequencing與其他定序研究方法中，所找到在國人樣本中具有多型性（polymorphism）之SNP點。 ● 其他與藥物反應、藥物代謝相關，如MHC，PGX等基因上的SNP點。	● TWB Array，共104,909個SNP點。 ● 根據TWB Array的使用經驗及Taiwan Biobank 1000名參與者次世代定序的結果，設計446K個SNP點。 ● 與臨床顯著相關的SNP點約105K個。 ● Thermo Fisher Scientific歷年替各國Biobank設計的疾病相關標記SNP。 ● 其他與藥物反應、藥物代謝相關，如MHC，PGX等基因上的SNP點。 ● 偵測拷貝數變異（CNV）的SNP點。

Taiwan View



Ref. version: [GRCh38]
Source: Healthy controls
Type and platform: [Next Generation Sequencing: Proton]
Number of subject: 517

Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y

chr1 **From** 1 **To** 249250621 GO

1 2 3 4 5 6 > Total:3825475

Chr	Pos	ID	Ref	Individuals	>>	Freq.	Other-platform	Anno.	LD
chr1	10108	rs62651026	C	41	>>	C:0.9878 T:0.0122		show	
chr1	10156		C	297	>>	C:0.9966 G:0.0034		show	
chr1	10162		C	428	>>	C:0.9766 G:0.0234		show	
chr1	10166		C	221	>>	C:0.9932 G:0.0068		show	
chr1	10168		C	464	>>	C:0.9935 G:0.0065		show	
chr1	10171		ACCC	371	>>	ACCC:0.9960 AGCC:0.0027 ACCG:0.0013		show	
chr1	10172		C	490	>>	C:0.9633 G:0.0367		show	
chr1	10231	rs775928745	C	356	>>	C:0.9944 A:0.0056		show	
chr1	10235	rs540431307	T	255	>>	T:0.9961 A:0.0020 C:0.0020		show	
chr1	10241	rs960927773	T	255	>>	T:0.9961 C:0.0039		show	
chr1	10248	rs148908337	A	101	>>	A:0.9950 T:0.0050		show	
chr1	10250	rs199706086	A	483	>>	A:0.9451 C:0.0549		show	
chr1	10257	rs111200574	A	465	>>	A:0.9742 C:0.0258		show	
chr1	10267	rs921020588	T	121	>>	T:0.9959 G:0.0041		show	
chr1	10390	rs766767872	C	99	>>	C:0.9949 A:0.0051		show	
chr1	10393		C	99	>>	C:0.9949 T:0.0051		show	

臺灣人體生物資料庫介紹(1)

https://www.youtube.com/watch?v=a4yK1AhUrRw

臺灣人體生物資料庫介紹(2)

https://www.youtube.com/watch?v=LmHJ1KFlW1s&ab_channel=TaiwanBiobank

國衛院推動癌症精準醫療！台灣禮來加入生物資料庫整合平台

作者 姚 惠茹 | 發布日期 2022 年 05 月 09 日 14:32 | 分類 生物科技 , 醫療科技

【NExT Forum】台大醫教授楊泮池：兩大關鍵助台建構精準醫療的大數據永續平台

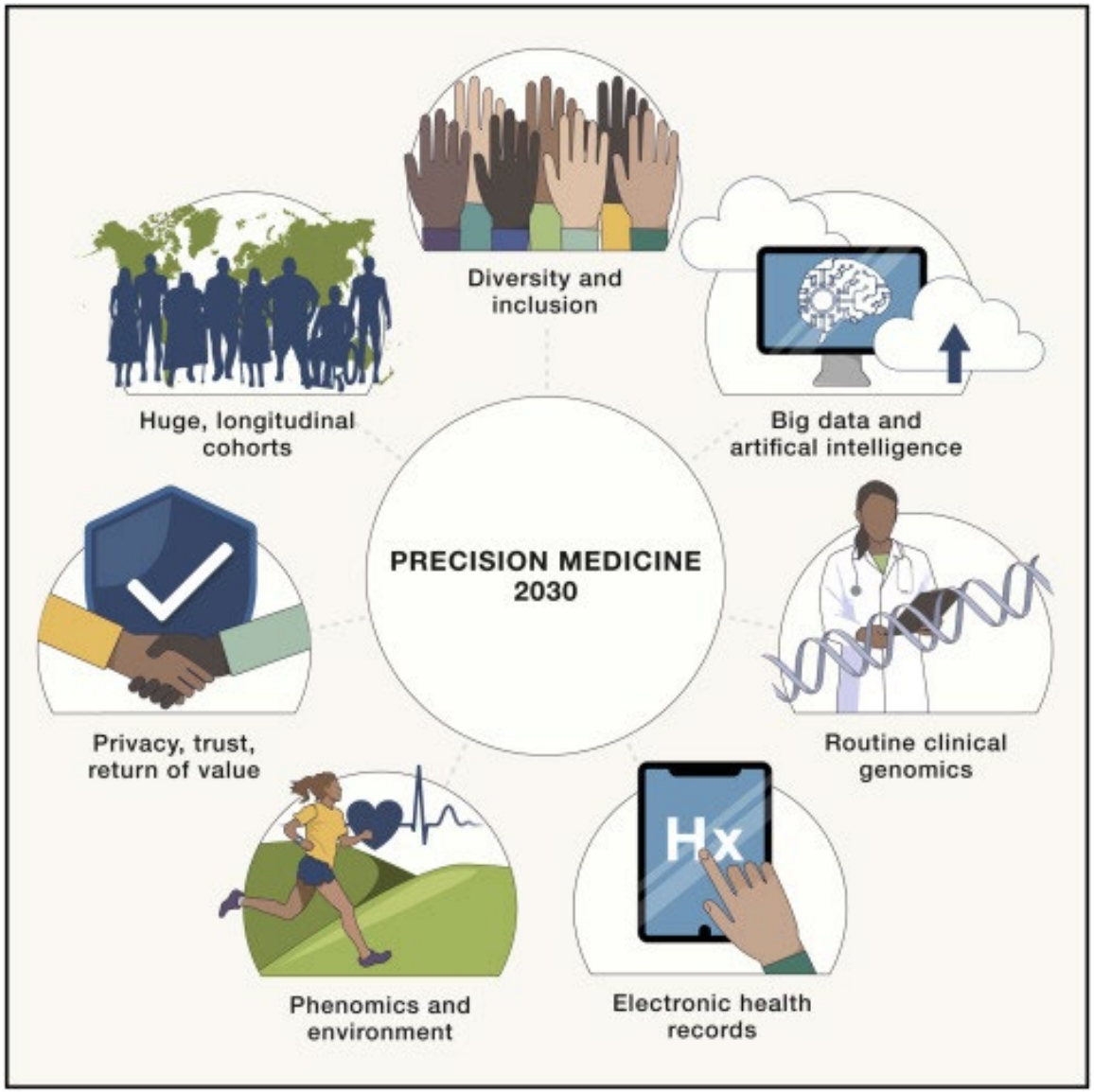
TPMI計畫破50萬人參與 增進台灣精準醫療量能

中央通訊社

2022/8/18 12:37（8/18 13:51 更新）

HTTPS://WWW.CNA.COM.TW/NEWS/AHEL/202208180109.ASPX

2030年的精準醫療



Denny and Collins (2021, *Cell*)

基因體協助精準診斷與醫療

http://www.hugo-international.org/blog/?p=117

http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/newssummary/news_5-7-2011-9-44-25

醫學影像協助疾病診斷與智慧健康

https://blog.radiology.virginia.edu/different-imaging-tests-explained/

生醫演進	大眾化醫療	個人化醫療	精準醫療	精準健康
	2002	2009	2013	2017 – 2021
推動政策	兩兆雙星產業發展計畫	臺灣生技起飛鑽石行動方案	臺灣生技產業起飛行動方案	5+2產業創新：生醫產業創新、臺灣精準健康

跨部會協力打造精準健康未來

扶植精準健康產業鏈

完善精準健康生態系

接軌國際佈局全球

行政院農業委員會 COUNCIL OF AGRICULTURE, EXECUTIVE YUAN
國家發展委員會 NATIONAL DEVELOPMENT COUNCIL
NSTC 國家科學及技術委員會 National Science and Technology Council
經濟部
金融監督管理委員會 Financial Supervisory Commission R.O.C. (Taiwan)

融合生醫跨域環境	整合串連園區聚落	強化國際鏈結
布局下世代人才	優化法規環境	完善健康大數據系統建置	建構友善投資環境

尾聲：

二十一世紀是精準和智慧健康的時代，也是大數據統計資料科學的黃金時代！

謝謝聆聽