Biological sample collection and study design for stroke genetics studies

2015/01/28

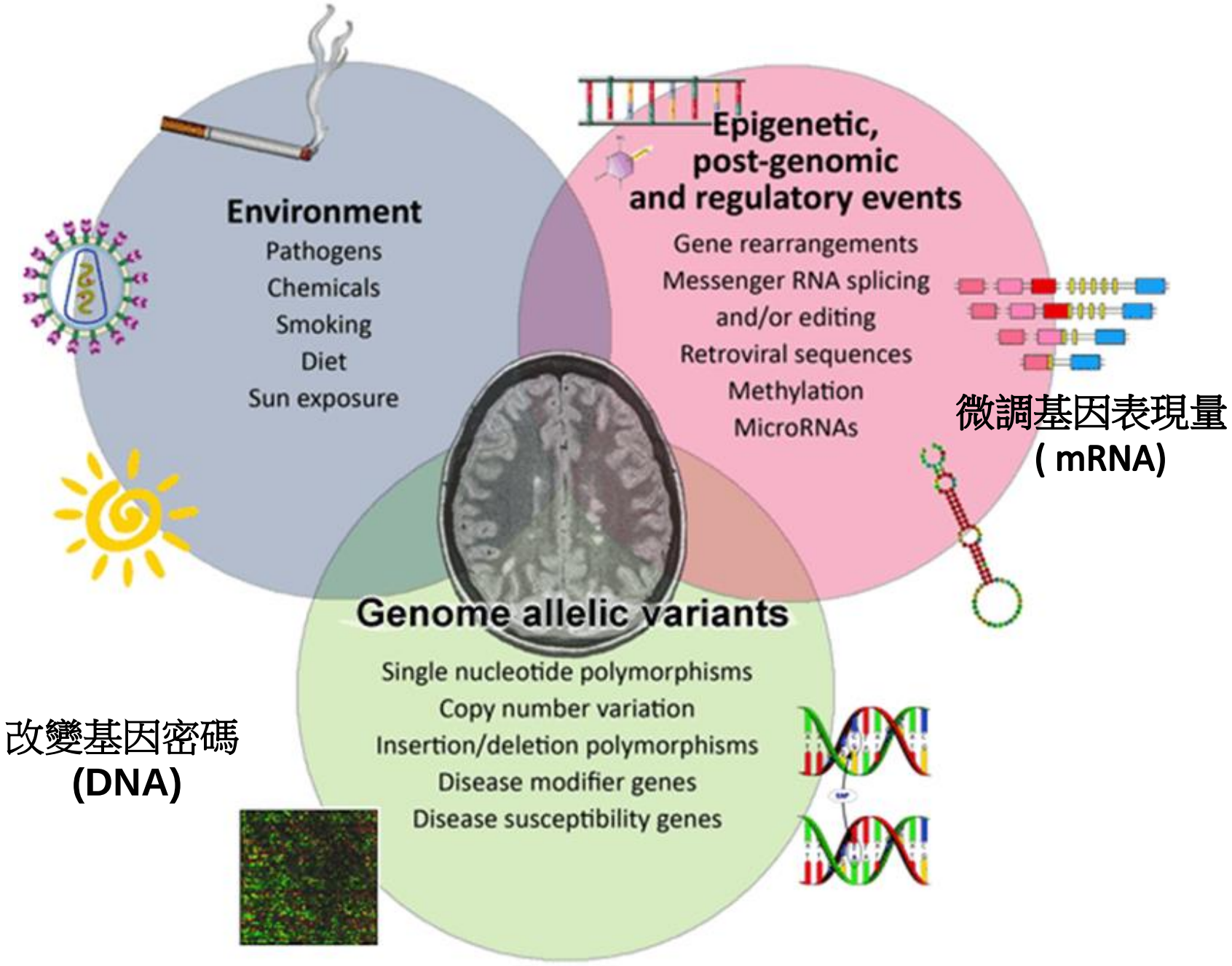
廖翊筑

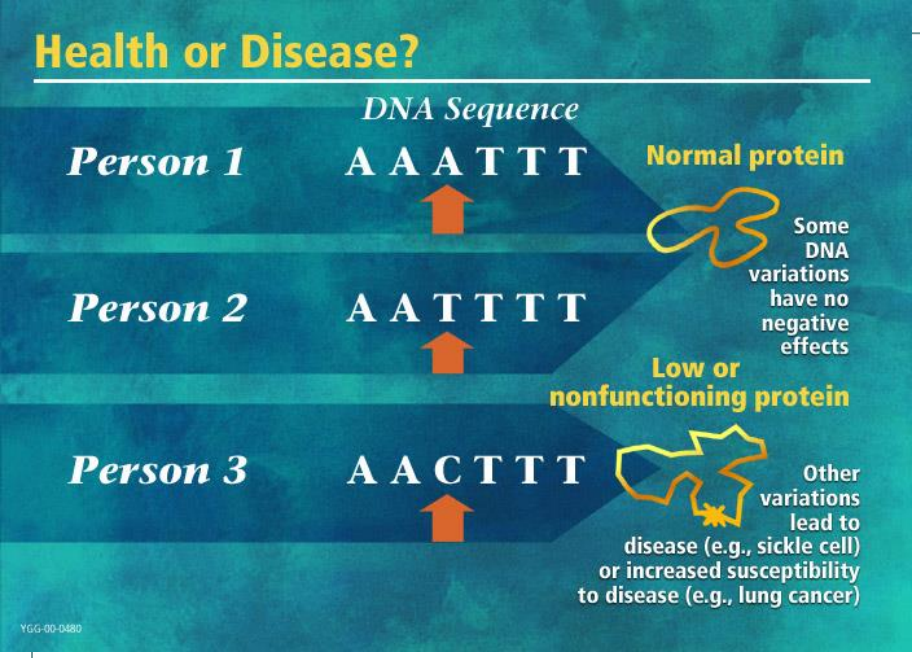
Outline

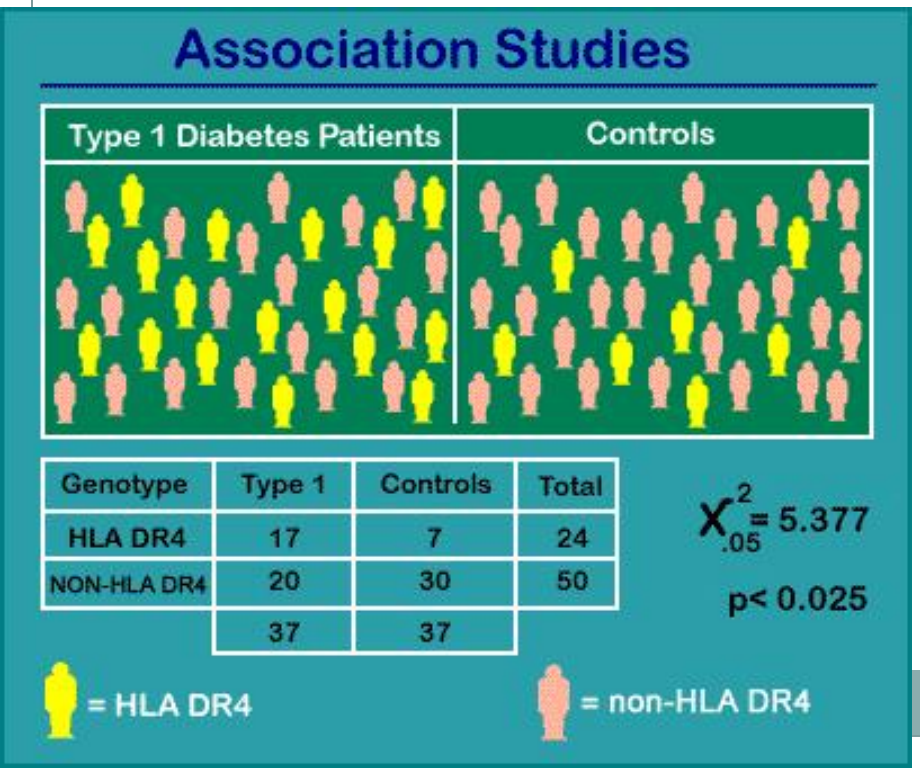
- **Stroke genetic studies**
  -- SNP association study
  -- GWAS
  -- Pharmaco-genetic studies
  -- Epigenetics study (DNA methylation, microRNAs)

- **What kinds of biological samples to be collected?**
  -- DNA, plasma/serum, lymphocyte

- **How to collect biological samples?**


Environment
Pathogens
Chemicals
Smoking
Diet
Sun exposure
Epigenetic, post-genomic and regulatory events
Gene rearrangements
Messenger RNA splicing and/or editing
Retroviral sequences
Methylation
MicroRNAs
Genome allelic variants
Single nucleotide polymorphisms
Copy number variation
Insertion/deletion polymorphisms
Disease modifier genes
Disease susceptibility genes
微調基因表現量
(mRNA)
改變基因密碼
(DNA)





SNP (single nucleotide polymorphism)

(1) 1 SNP every 100-300 bp
(2) As of October 2014, dbSNP listed 112,736,879 SNPs
(3) < 1% mutation
> 1% SNP

Genetic association study

ApoE ε2/ε3/ε4 vs. AD risks

145 genes in 8 functional domains average 4.8 SNPs per gene (range 3-15)

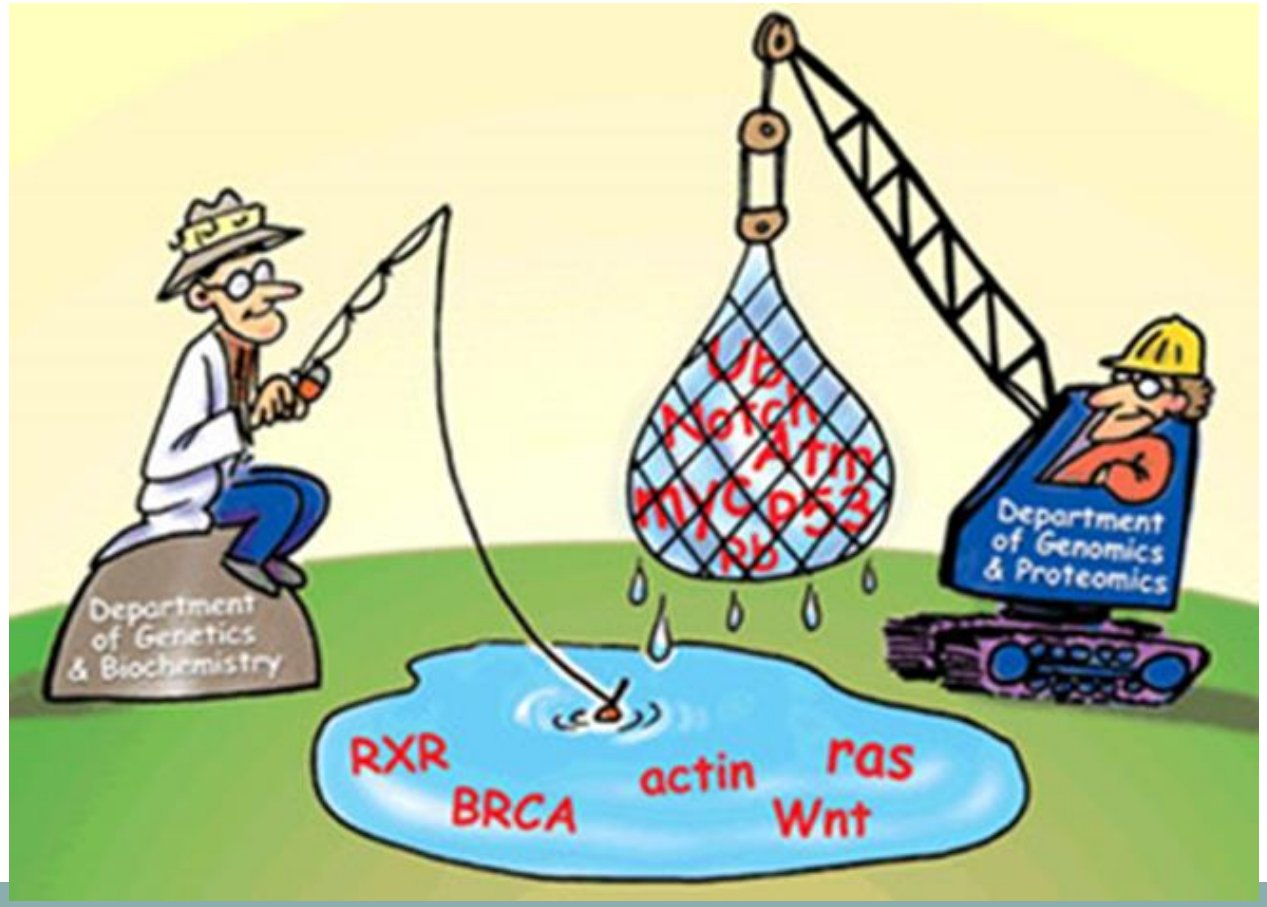
Liao et al. Stroke. 2008 Dec;39(12):3159-65.

Table I. Information on 145 Candidate Genes Analyzed in This Study

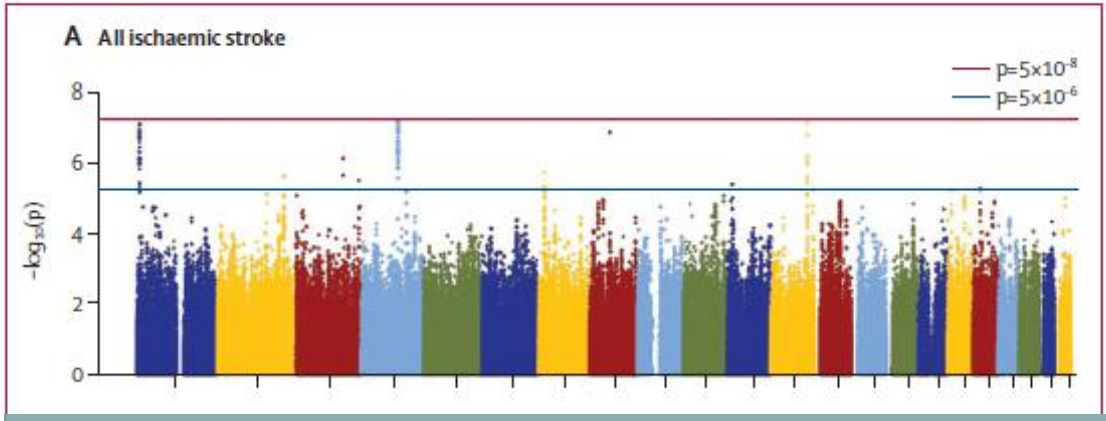
Category	N	Genes*
Extracellular matrix	15	*ANXA5, DCN, GJA4, LMNA, MGP, MMP1, MMP12, MMP2, MMP3, MMP9, TIMP1, TIMP2, TGFB1, TGFB2, TGFB3*
Hemostasis	21	*CD36, CPB2, FGA, FGB, FGG, FGL2, FSBP, GP1BA, PDGFB, PLAT, PLAU, PLAUR, PLG, SERPINB2, SERPINC1, SERPINE1, THPO, HFE, VWF, THBS1, THBS4,*
Endothelium function	24	*CD9, HABP2, ICAM1, ITGA2, ITGA2B, ITGA8, ITGB3, SELE, SELL, SELP, SELPLG, VCAM1, CBS, DES, EDN1, MTHFR, MTR, MTRR, NNMT, NOS1, NOS2A, NOS3, PDE4D, VEGF*
Renin–angiotensin system	15	*ACE, ACE2, AGT, AGTR1, AGTR2, ADD1, CMA1, NPPA, NPR1, NPY, REN, SCNN1A, SCNN1B, SCNN1G, TH*
Inflammation	24	*CCL2, CCR2, CEBPB, CKLF, CD14, CRP, CXCL12, IL1RN, IL10, IL1A, IL1B, IL6, IL6R, IL8, LTA, LTB, PTGS2, TFP1, (TGFB1, TGFB2, TGFB3*), TLR4, MHC2TA, THBD, TNF, TNFSF4, BDKRB2*
Antioxidation	5	*CYBA, PON1, SOD2, SOD3, XRCC1*
Glucose metabolism	21	*ENPP1, GAD2, GCK, IGF1, IGF2, IGF2AS, INS, INSR, IRS1, pck1, PPARA, PPARG, ADIPOQ, ADRA1A, ADRA2A, ADRB1, ADRB2, ADRB3, GNB3, LEP, LEPR*
Lipid metabolism	20	*ABCA1, ALOX5, ALOX5AP, APOA5, APOB, APOE, CETP, LPL, CYP11B2, CYP7A1, ESR1, FABP2, LIPC, LIPE, LIPG, LRP1, MTP, OLR1, PLTP, SCARB1*

*Genes with pleiotropic effect are listed in multiple categories.

- **GWAS (genome-wide association study) :**
  Step 1: Identifies novel genes
  Step 2: validates its biological impact on diseases



	Number of cases	Number of CS cases	Number of LVD cases	Number of SVD cases	Number of controls	Ancestry	Study design	Genotyping
Discovery cohorts								
ARIC	385	93	31	63	8803	European	Population-based	Affymetrix 6.0
ASGC	1162	240	421	310	1244	European	Cross-sectional	Illumina 610
BRAINS	361	29	120	97	444	European	Cross-sectional	Illumina 660
CHS	454	147	..	73	2817	European	Population-based	Illumina 370
deCODE	2391	399	255	240	26 970	European	Cross-sectional	Illumina 317/370
FHS	171	48	..	..	4164	European	Population-based	Affymetrix 550
GEOS	448	90	37	54	498	European	Cross-sectional	Illumina HumanOmni1
HPS	578	..	..	..	468	European	Cross-sectional	Illumina 610
HVH	566	88	61	173	1290	European	Cross-sectional	Illumina 370
ISGS/SWISS	1070	247	229	201	2329	European	Cross-sectional	Illumina 550/610/660
MGH-GASROS	516	169	95	38	1202	European	Cross-sectional	Affymetrix 6.0
Milano	372	25	74	65	407	European	Cross-sectional	Illumina 610/660
Rotterdam	367	..	..	..	5396	European	Population-based	Illumina 550
WTCCC2-Munich	1174	330	346	106	797	European	Cross-sectional	Illumina 660
WTCCC2-UK	2374	460	498	474	5175	European	Cross-sectional	Illumina 660
Total (discovery)	12 389	2365	2167	1894	62 004	..	..	..



0.5-2.5 million SNPs genotyped
→ imputation to 5-15 million SNPs

Bonferroni correction
→ 1 million SNPs, 0.05 /10^{-6}

Lancet Neurol 2012:11:951-62

Replication cohorts								
Barcelona	439	179	110	150	404	European	Cross-sectional	Sequenom
BSS	225	11	93	90	312	European	Cross-sectional	Sequenom
Copenhagen	730	..	..	..	1545	European	Cross-sectional	TaqMan
ESS	276	40	20	69	940	European	Cross-sectional	TaqMan/Illumina 610
Glasgow	675	125	91	150	940	European	Cross-sectional	Sequenom/Illumina 610
Go-Darts*	737	130	259	..	8424	European	Cross-sectional	Affymetrix 6.0/Illumina Cardio-metabochip
Graz	657	116	108	207	848	European	Cross-sectional	Sequenom/Illumina 610
Interstroke*	872	143	198	238	926	European	Cross-sectional	Illumina Cardio-metabochip
Krakow	1235	377	152	171	584	European	Cross-sectional	Sequenom
Leuven	458	195	83	63	391	European	Cross-sectional	Sequenom
Lund	424	140	21	94	466	European	Cross-sectional	Sequenom
Munster	1232	478	528	224	1053	European	Cross-sectional	Sequenom
Portugal	539	..	..	..	507	European	Cross-sectional	Sequenom
RACE (Pakistan)*	1322	225	195	189	1143	Pakistani	Cross-sectional	Illumina 660
SMART	623	30	368	195	6712	European	Population-based	Sequenom
Sweden	876	157	177	75	742	European	Cross-sectional	Sequenom
VISP*	1725	..	..	..	1047	European	Cross-sectional	Illumina HumanOmni1
WHI*	302	42	31	78	2099	European	Population-based	Illumina Omni-Quad
Total (replication)	13347	2388	2434	1993	29083	..	..	..

Susceptible genes identified by GWAS: meta-analysis

	Reference	Stroke subtype*	Gene	SNP with strongest association	Risk allele	Risk allele frequency (%)†	Odds ratio (95% CI)	p
Ischaemic stroke‡								
9q34	10,58	Cardioembolic; large vessel	*ABO*	rs505922	C	39%	1·13 (1·11–1·15); 1·23 (1·07–1·18)	<0·0001; <0·0001
4q25	10,59–62	Cardioembolic	*PITX2*	rs6843082	G	21%	1·36 (1·27–1·47)	$2·8 \times 10^{-16}$
16q22	10,63	Cardioembolic	*ZFHX3*	rs879324	A	19%	1·25 (1·15–1·35)	$2·3 \times 10^{-8}$
7p21	10,64	Large vessel	*HDAC9*	rs2107595	A	16%	1·39 (1·27–1·53)	$2·0 \times 10^{-16}$
6p21	10,65	Large vessel	*SUPT3H/CDC5L*	rs556621	A	33%	1·21 (1·12–1·28)	$4·7 \times 10^{-8}$
9p21	10	Large vessel	*CDKN2A/CDKN2B*	rs2383207	G	52%	1·15 (1·08–1·23)	$3·3 \times 10^{-5}$
Intracerebral haemorrhage								
1q22	66	Non-lobar ICH	*PMF1/SLC25A44*	rs2984613	C	32%	1·33 (1·22–1·46)	$2·2 \times 10^{-10}$
19q13	67	Lobar ICH	*APOE*	rs429358/rs7412	ε2	7%	1·82 (1·50–2·23)	$6·6 \times 10^{-10}$
19q13	67	Lobar ICH	*APOE*	rs429358/rs7412	ε4	12%	2·20 (1·85–2·63)	$2·4 \times 10^{-11}$
13q34	68	Lobar and non-lobar ICH	*COL4A1*	1055C→T; 1612C→G	T; G	<1%	..	..
13q34	69	Lobar and non-lobar ICH	*COL4A2*	3448C→A; 5068G→A; 3368A→G	A; A; G	<1%	..	..

Af

CVD, MI

PITX2: paired-like homeodomain 2; ZFHX3: zinc finger homeobox 3; HDAC9: histone deacetylase 9
CDKN2A: cyclin-dependent kinase inhibitor 2A; COL4A1: collagen, type IV, alpha 1

Lancet Neurol 2014;13:405-18

Polygenic risk score for disease prediction, personalized medicine

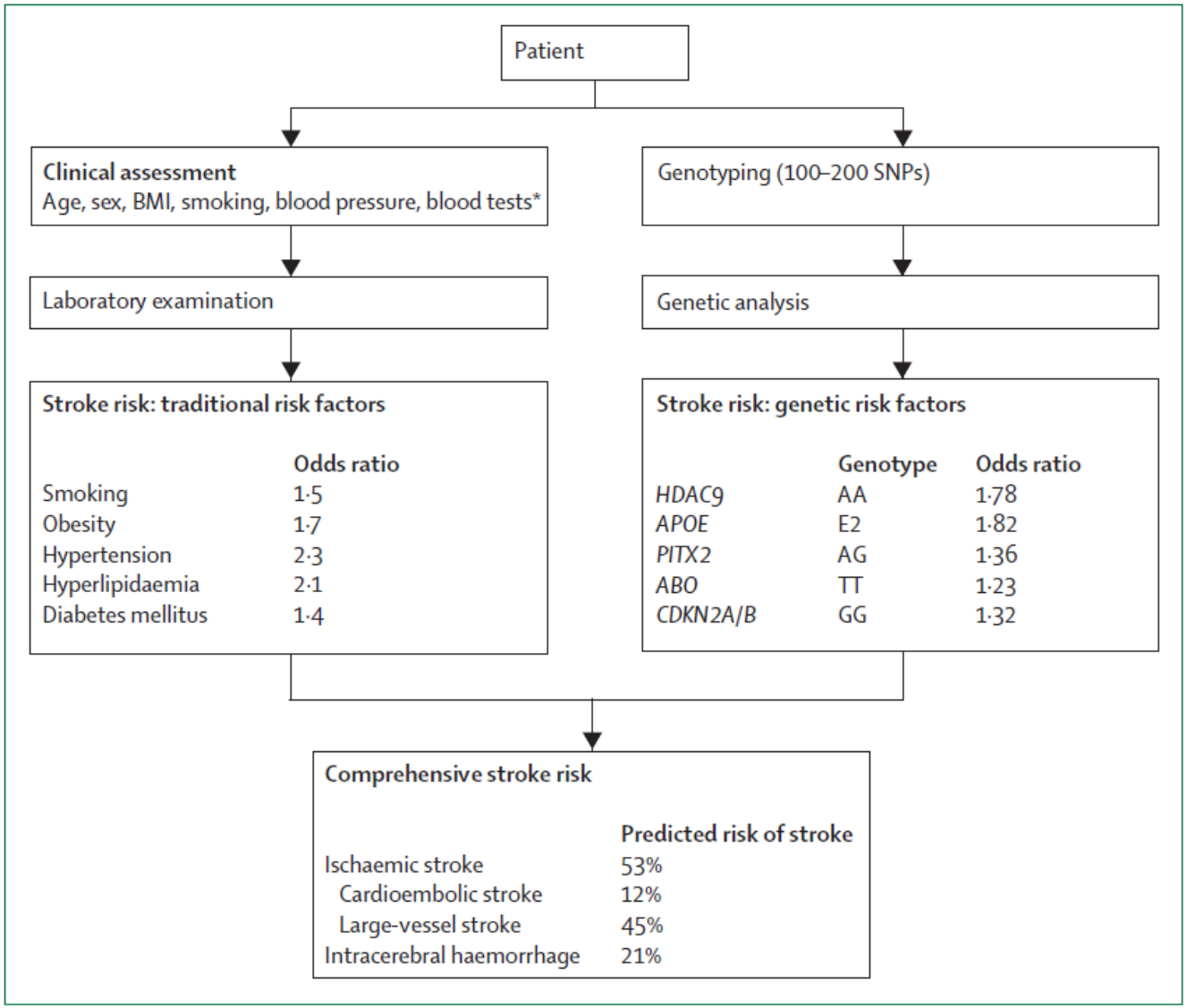
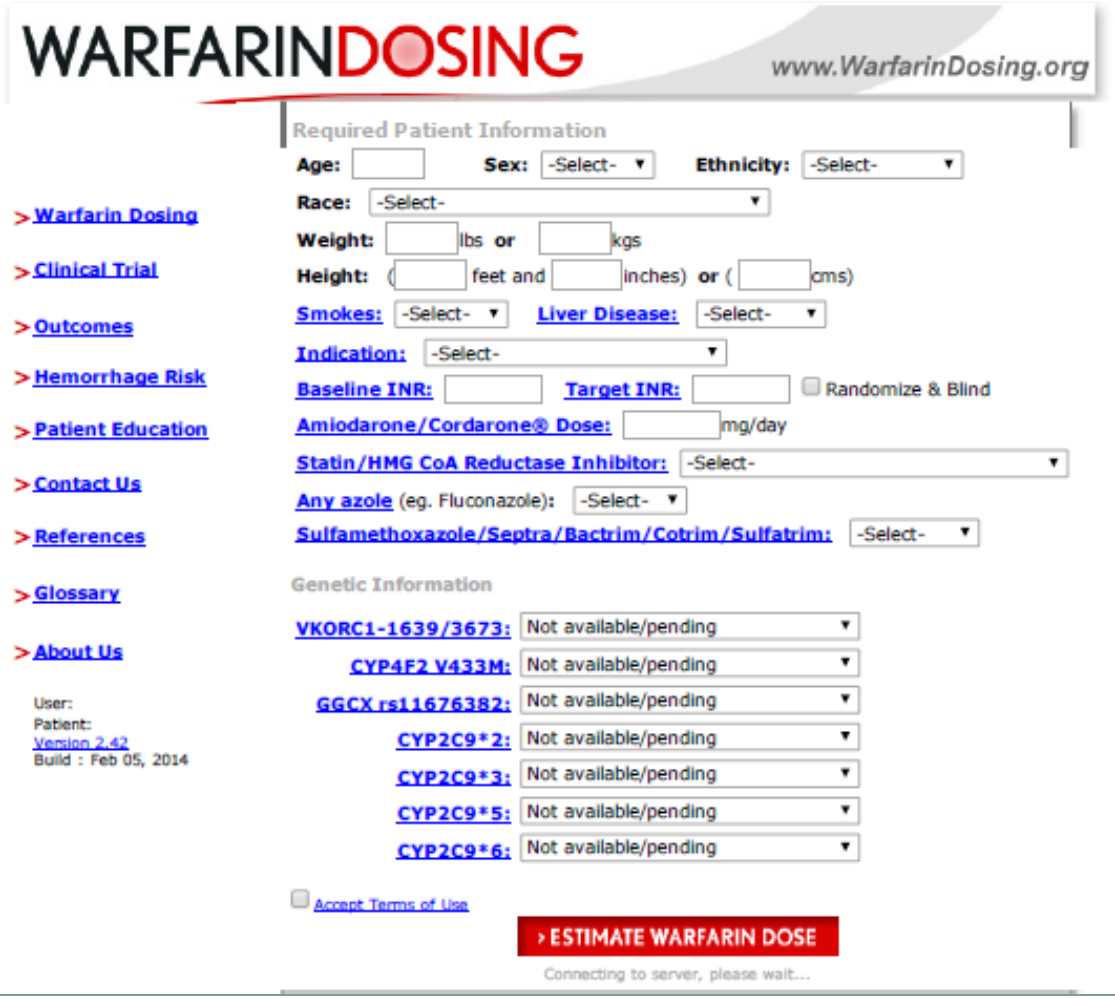


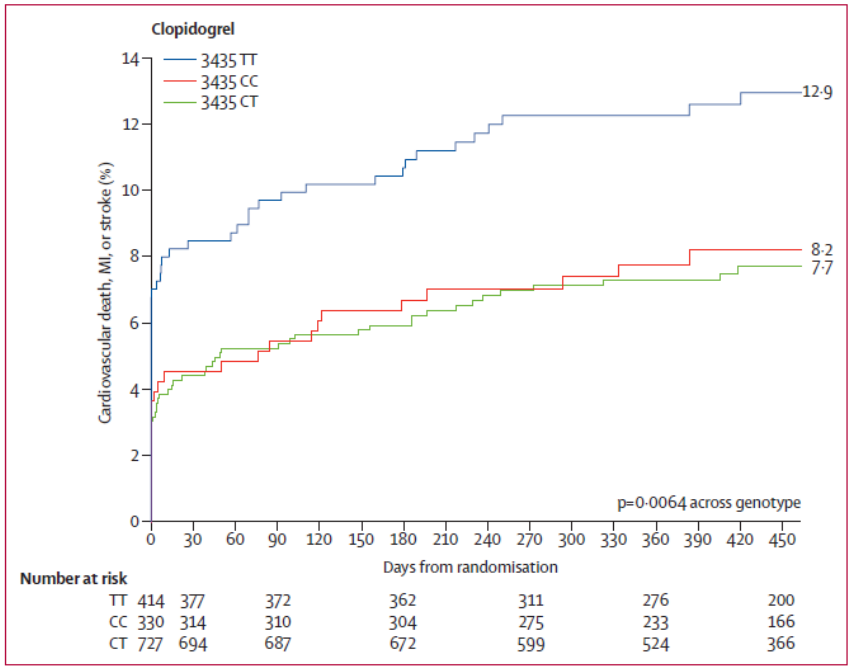
Figure 2: Example of personalised medicine in stroke

Lancet Neurol 2014;13:405-18

Warfarin dosing according to CYP2C9/VKORC1 (Vit K epoxide reductase complex 1) polymorphism

35% of variability in response to warfarin is related to VKORC1 & CYP2C9





CYP2C19*2 allele reduced clopidogrel conc. by 30%, increase risk of CVD

CES1 rs2244613 reduced dabigatran conc. by 15%, reduced bleeding risk by 33%

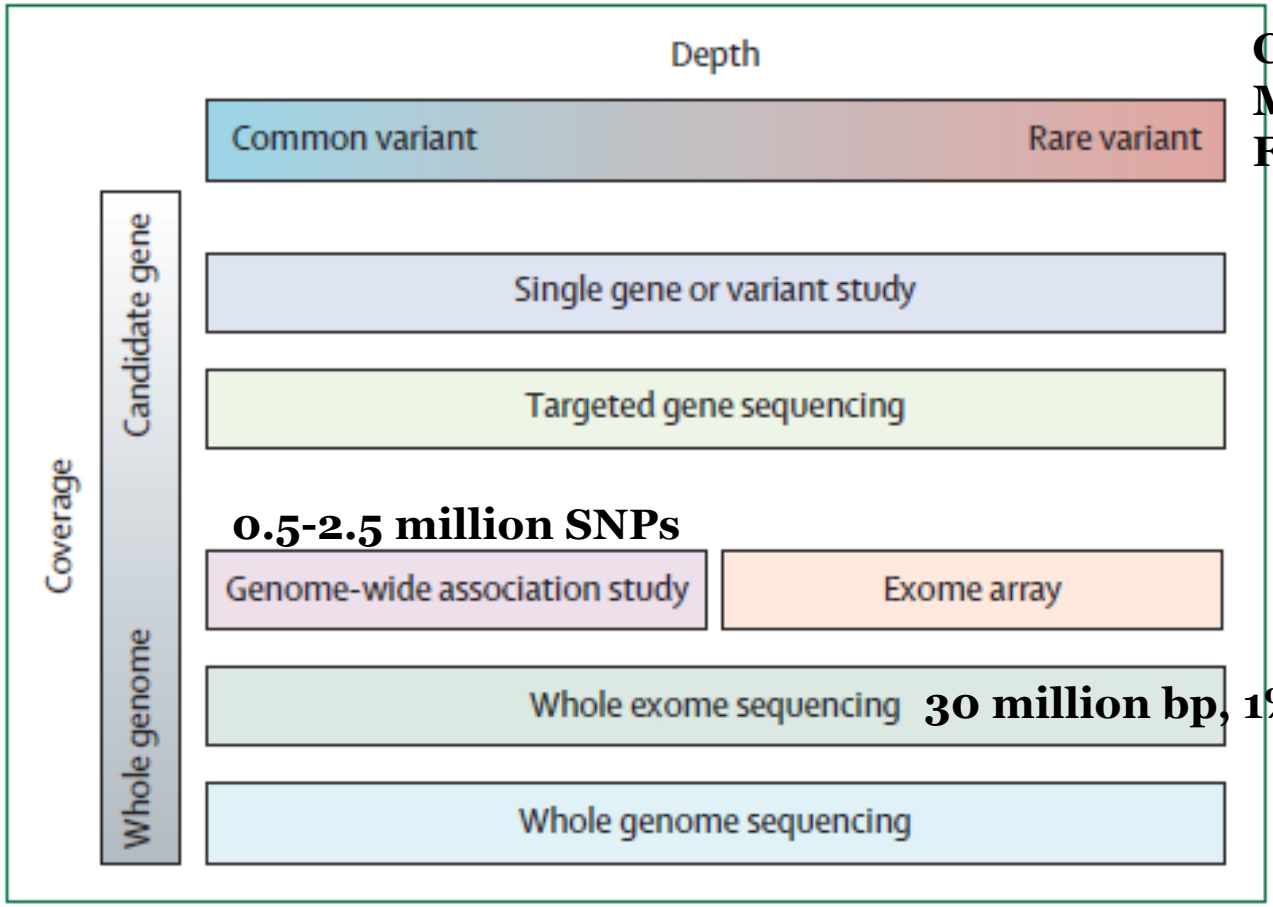
Lancet 2010:376:1312-9

Table 7. Summary of 3 Recent Meta-Analyses: *CYP2c19* Variants and Events

Reference	Population	Outcome	Studies, n	Patients, n	Events, n	1 or 2 vs 0, RR (95% CI) for Adverse Cardiovascular Events
Mega et al[228]	Invasively managed	CVD death, MI, stroke	9	9685	863	1.55 (1.11–2.17)
Bauer et al[229]	Established coronary disease	Same	12	18529	1676	1.11 (0.89–1.39)
Holmes et al[230]	All	Same+death	32	42016	3545	1.18 (1.09–1.28)

CI Indicates confidence interval; CVD, cardiovascular disease; MI, myocardial infarction; and RR, relative risk.

Circulation 2013;128:2813-51



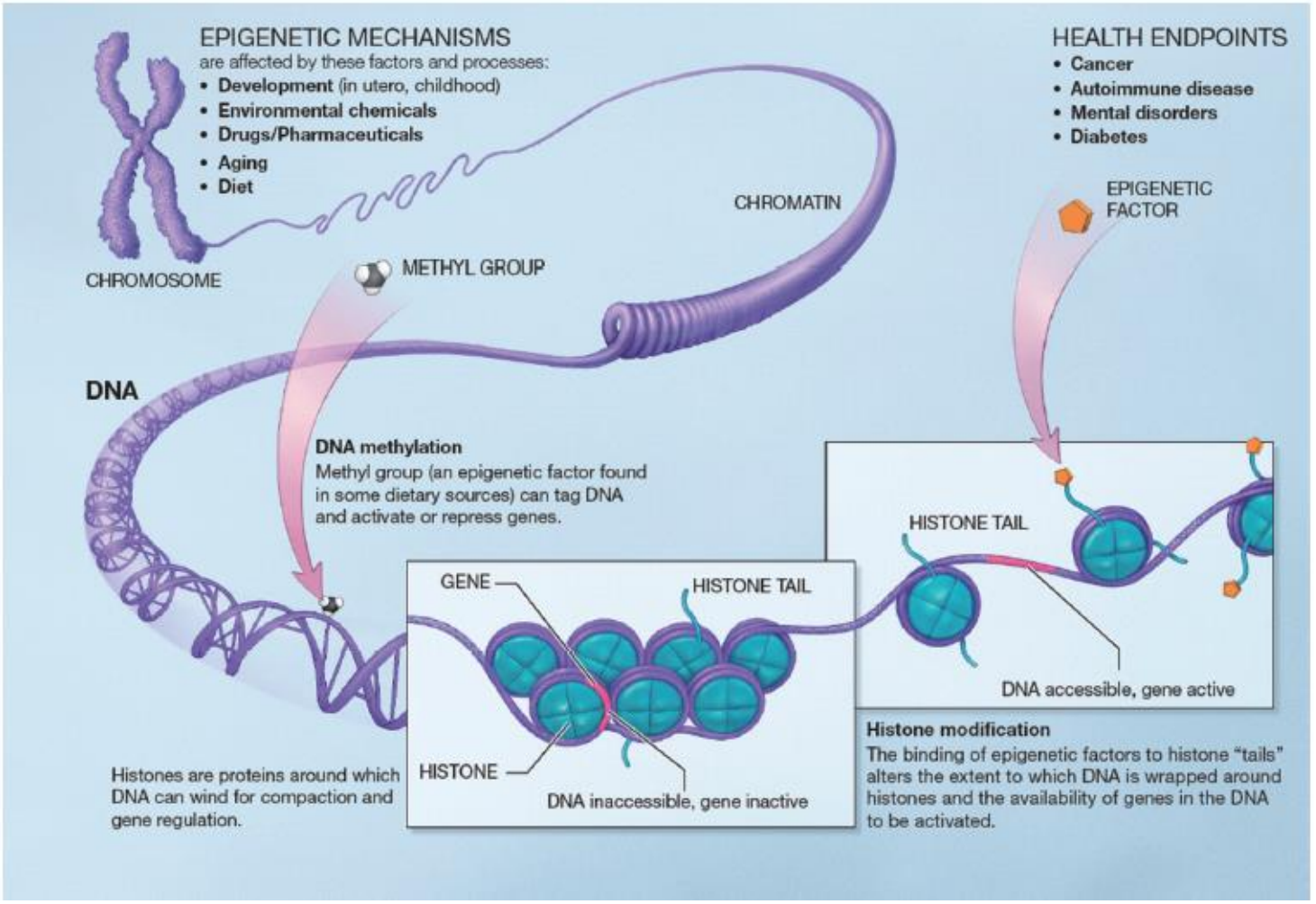
CADASIL, MELAS, Fabry's Diz

30 million bp, 1% of genome

Figure 1: Genotyping strategies

Lancet Neurol 2014;13:405-18

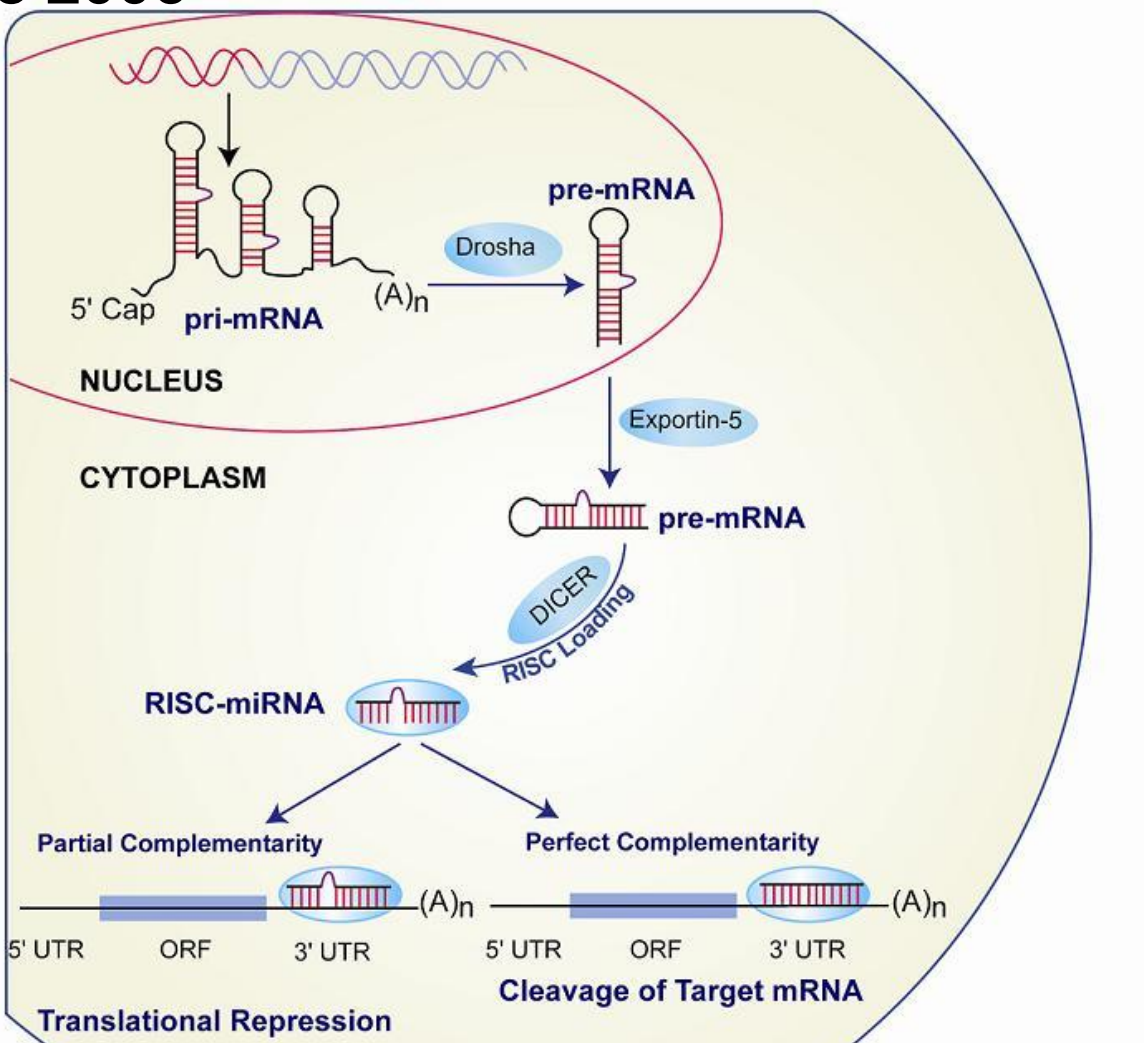
Epigenetic studies: methylation, histone modification, microRNA



The Nobel Prize in Physiology or Medicine 2006

Andrew Z. Fire

Craig C. Mello



miR, microRNA
(微小RNA)
抑制基因轉譯

siRNA, small interfering RNA
(小干擾RNA)
銷毀**mRNA**→ 基因沉默

Let-7g improved multiple endothelial functions through targeting TGF-b and SIRT-1 signaling

Liao et al. JACC 2014:63:1685-94

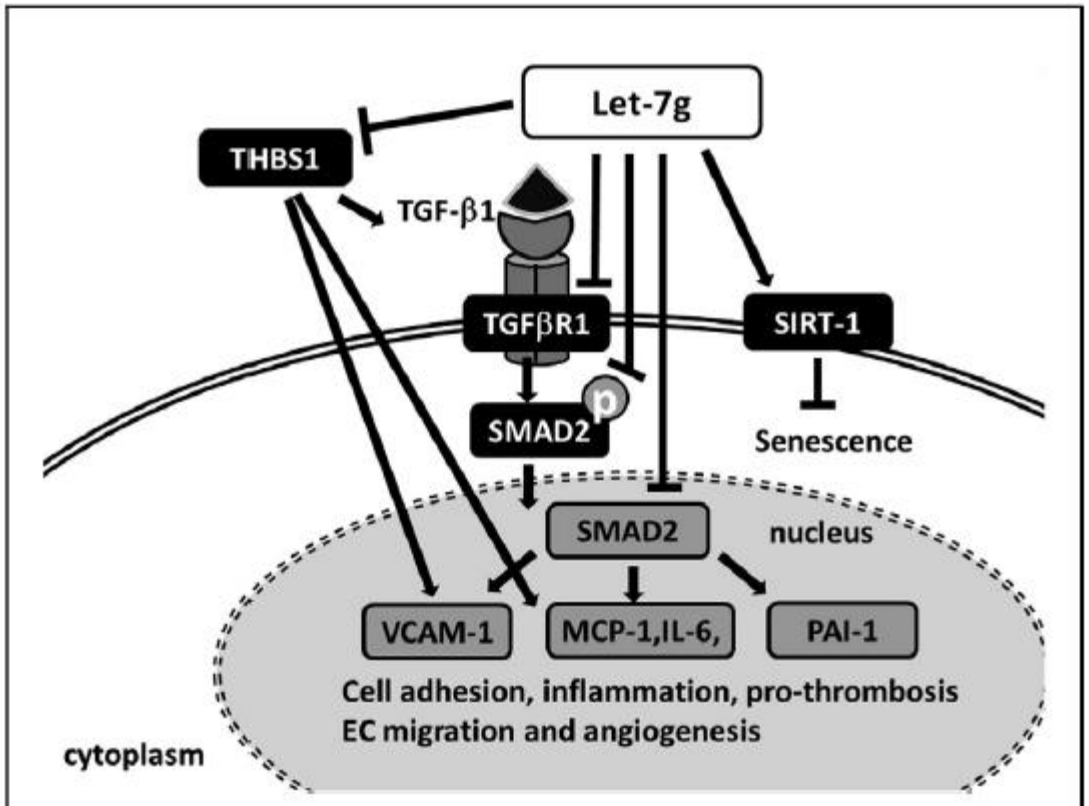
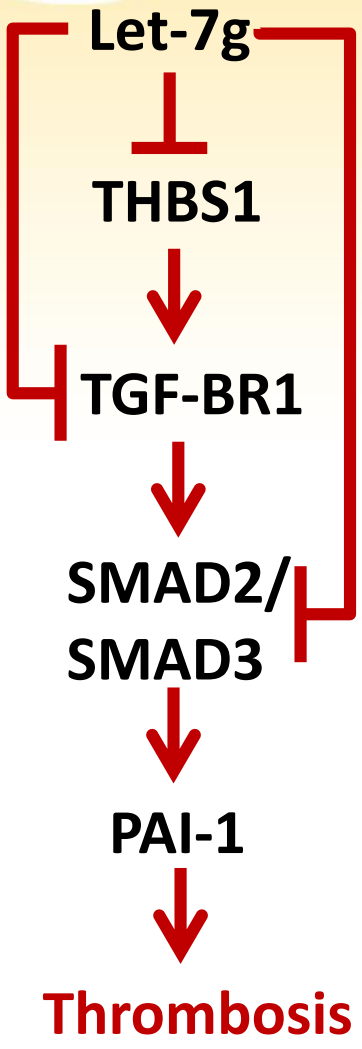
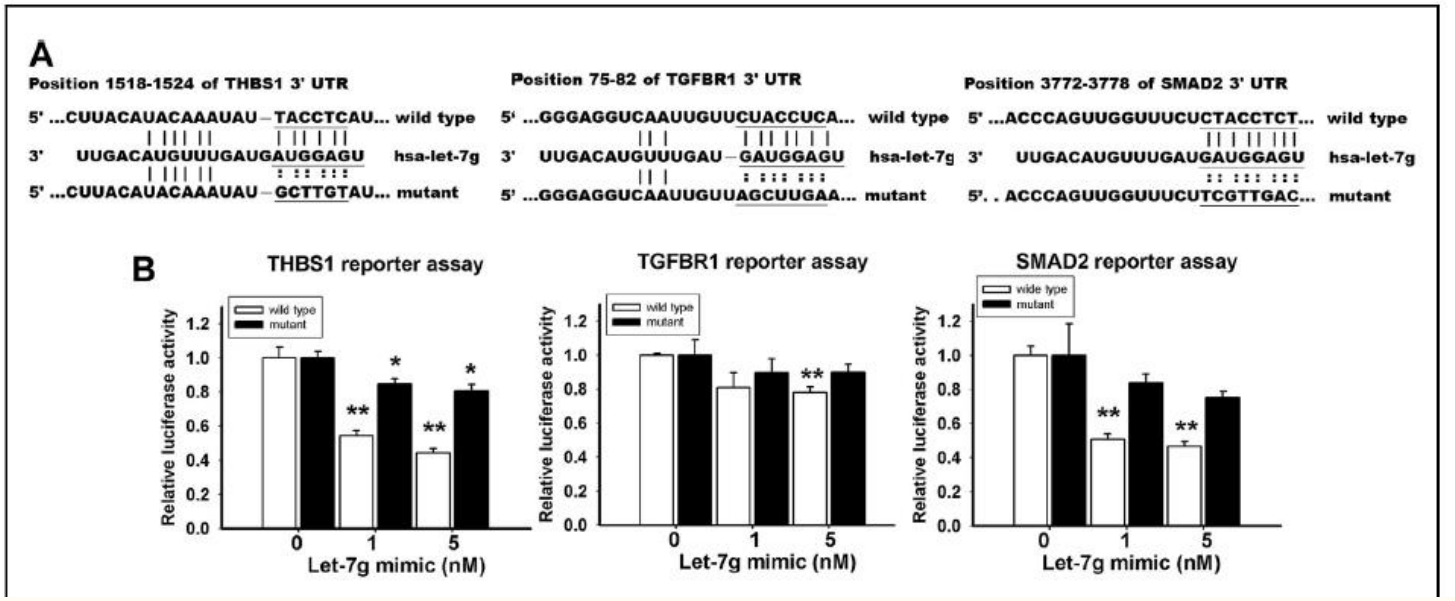


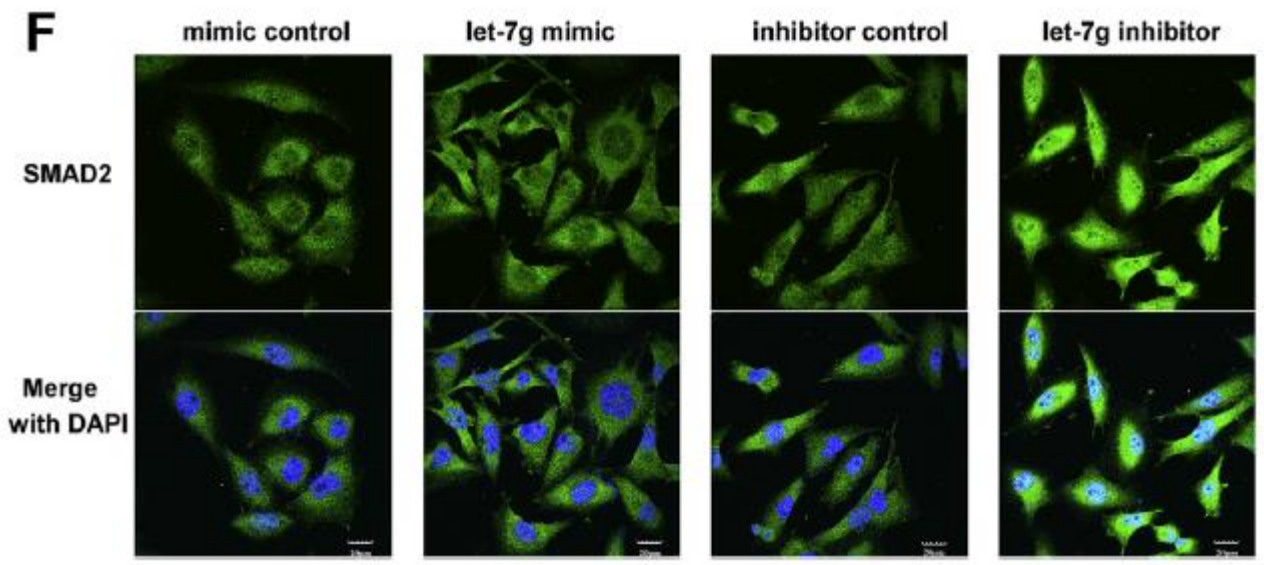
Figure 6 Schematic Shows Pleotropic Effects of Let-7g on Endothelial Functions

Abbreviations as in Figures 1 and 3.

Coagulation/ Thrombosis







Liao et al. JACC 2014:63:1685-94

Apo E KO mice HF diet

Apo E KO mice HF diet

inject lentivirus empty vector

Let-7g over-exp plasmids

per week x 12 weeks

inject lentivirus empty vector

Let-7g sponge plasmids

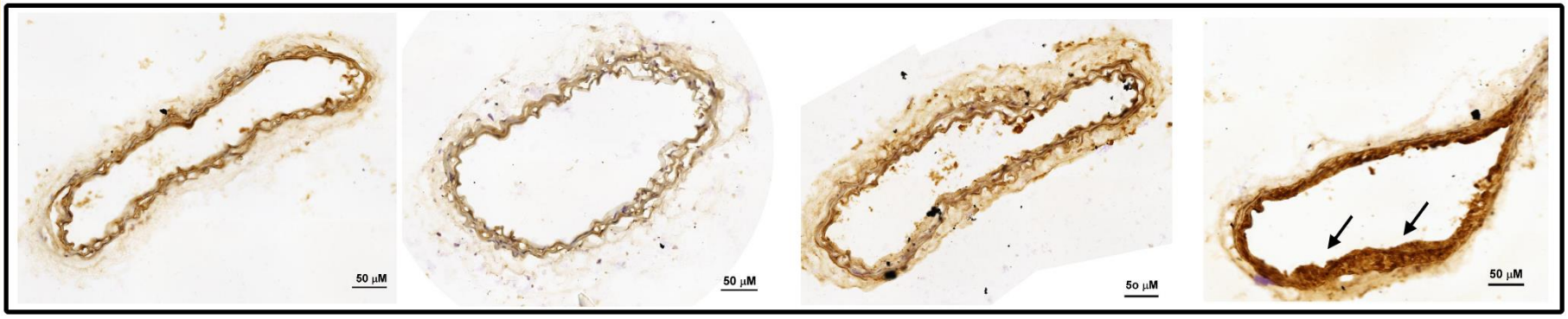
per week x 9 weeks

sacrifice mice and collect carotid arteries

動物實驗部分謹謝王永松博士, 鄭欣芸小姐, 林紋璉小姐, 蘇信吉先生協助進行

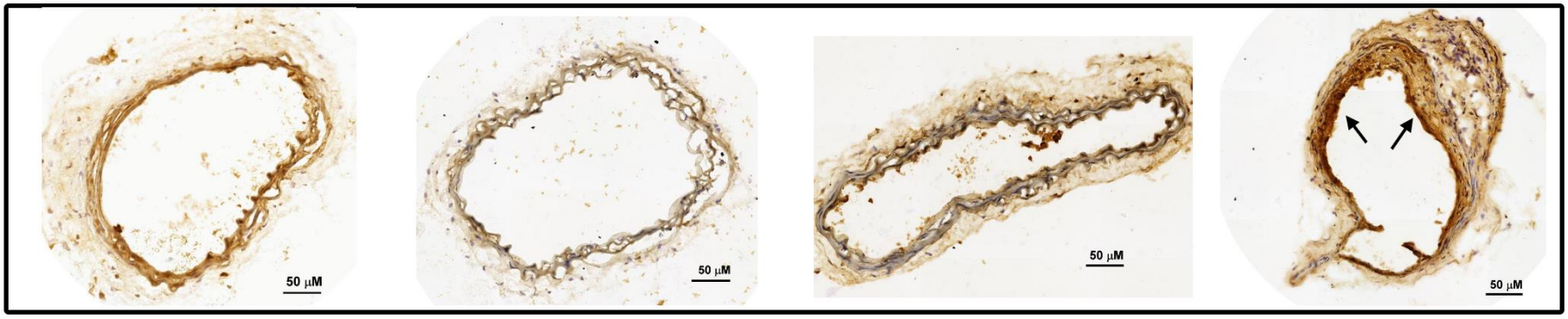
Let-7g inhibited PAI-1 expression in vivo:
IHC (免疫組織化學染色)stain in carotid artery of APOE-KO mice

PAI-1



HF 12 wk, empty vector | HF 12 wk, let-7g exp plasmids | HF 9 wk, empty vector | HF 9 wk, let-7g sponge

pSMAD2



HF 12 wk, empty vector | HF 12 wk, let-7g exp plasmids | HF 9 wk, empty vector | HF 9 wk, let-7g sponge

Validate Let-7g's effect on lacunar stroke patients

Coagulation/ Thrombosis

Vasodilatation

Plasma PAI-1 levels

Plasma ADMA levels (eNOS inhibitor)

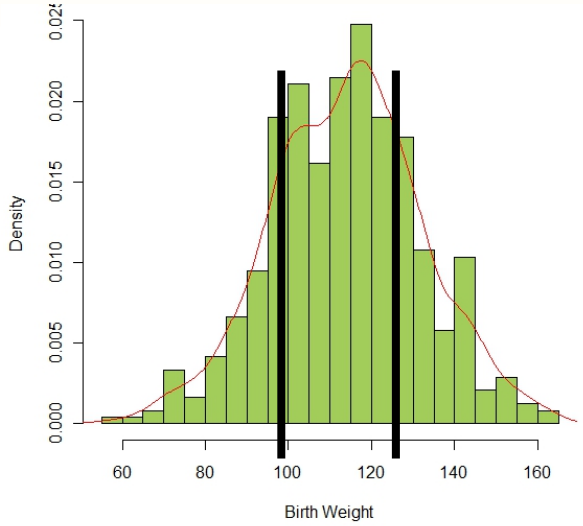
Correlated to serum Let-7g level ?

Study Design:

60 stroke patients (lacunar infarction)

→ Upper tertile or lower tertile of let-7g levels

→ measure their plasma levels of PAI-1 & ADMA



Higher PAI-1 expressions levels in lacunar stroke patients with low-let-7g levels

	Subjects with low let-7g levels	Subjects with high let-7g levels	Student's t test
Serum let-7g ratio ($2^{-\Delta\Delta Ct}$)*	0.48 ± 0.27 ↓	2.73 ± 1.36 ↑	$p < 0.001$
Plasma PAI-1 levels (ng/ml)	5.61 ± 3.24 ↑	3.70 ± 1.85 ↓	$p = 0.04$
Plasma ADMA (umol/L)	0.60 ± 0.19	0.58 ± 0.21	$p = 0.78$

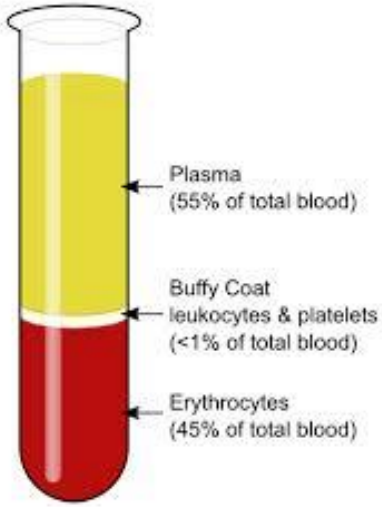
What you can get

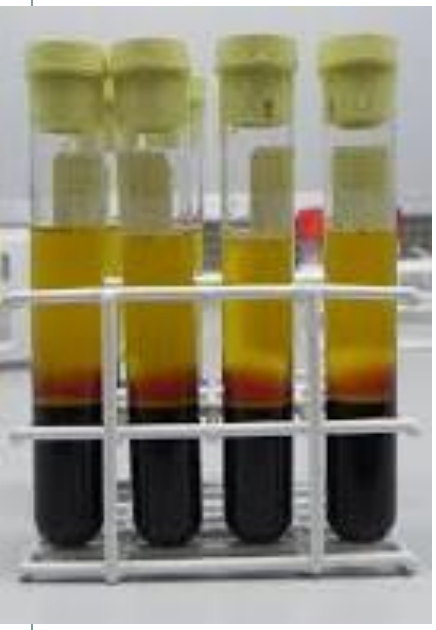
- ✓ DNA → for SNP, GWAS, DNA methylation/histone acetylation
- ✓ RNA (extracted from lymphocyte) → for gene expression levels, lymphocyte ≠ brain
- ✓ Lymphocyte → for culture or immortalization
- ✓ Plasma, Serum → cytokine (biomarkers, MCP-1, CRP, ICAM-1, VCAM-1, IL-6), microRNA

How to collect

- (1) **紫頭採血管** (含微量EDTA.K3，可以抑制血液凝集，EDTA可抑制DNases的活性，防止DNA受到DNases的作用)
  → 抽buffy coat 後可抽DNA
  → 離心後上清液是血漿 (plasma),建議加蛋白酶抑制劑, 分成小管保存為佳 (250-500 ul or 1 ml 裝)

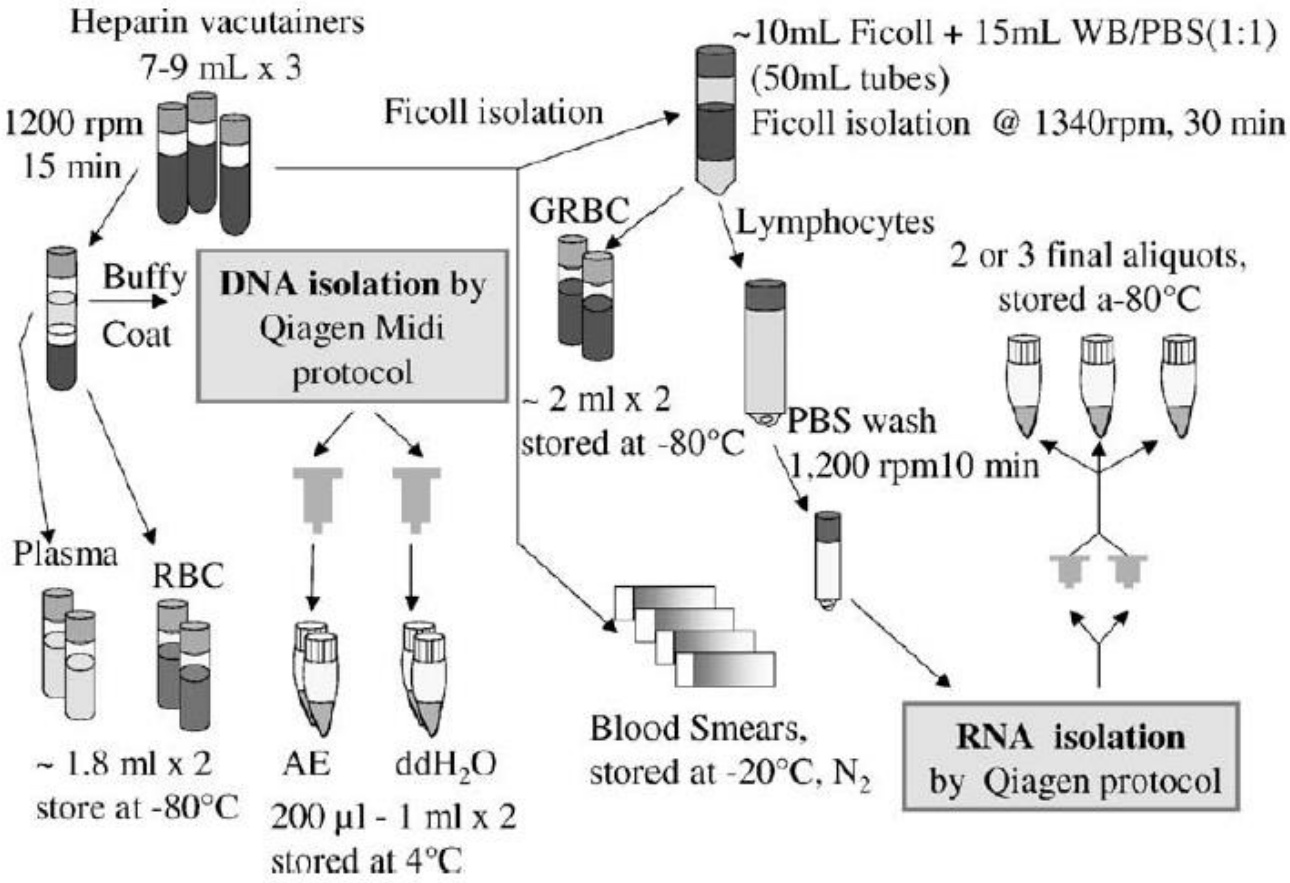
- (2) **紅頭核醫管**
  → 離心後上清液是血清 (serum),建議加蛋白酶抑制劑, 分成小管保存為佳 (250-500 ul or 1 ml 裝)
- 抽了血馬上處理, 過程都在冰上, 處理好後放-80 C 保存
- 抽血的日期要記錄(disease acute stage or not)
- IRB

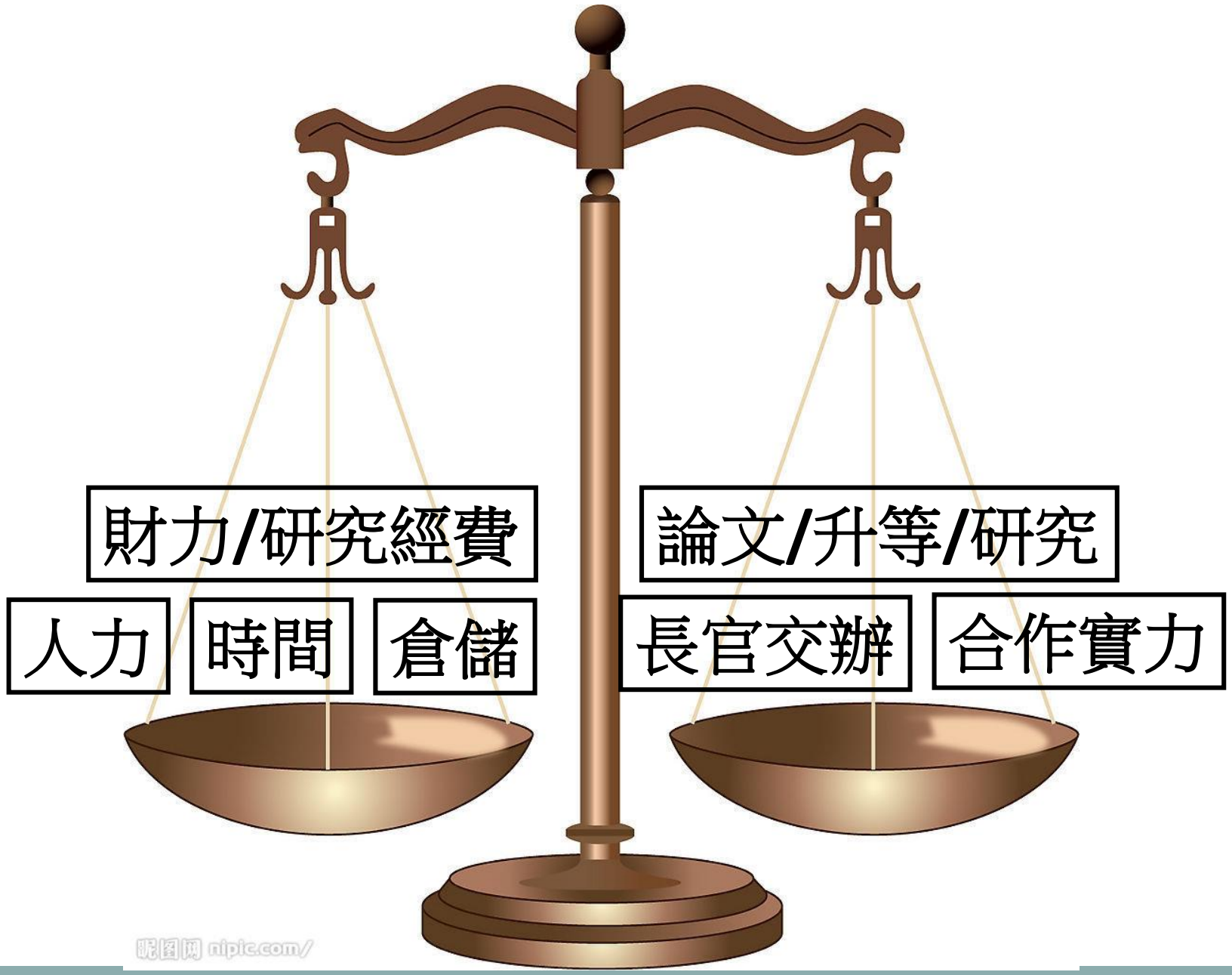




1. 抽lymphocyte & 抽 buffy coat/plasma 不能兩全

2. 要先抽成DNA 保存 或以buffy coat 形式保存?




財力/研究經費
人力
時間
倉儲
論文/升等/研究
長官交辦
合作實力

THANKS FOR ATTENTION !!

	Adult height (n of loci per 5000 individuals)	Crohn's disease (n of loci per 1000 cases and 1000 controls)	Schizophrenia (n of loci per 3000 cases and 3000 controls)	Ischaemic stroke (n of loci per 3000 cases and 10 000 controls)
1×	0	2	1	1
2×	2	4	2	2
3×	7	5	6	6
9×	68	51	62	NA
18×	180	NA	NA	NA

Increasing the total size of the sample (1×, 2×, etc) in genome-wide association studies yields consistently more susceptibility loci with genome-wide significance. The initial sample size for every trait differs, reflecting the fact that different sample sizes were needed to uncover the first genome-wide association locus. Data are measured values. NA=data not available (studies not done).

Table 1: **Loci with genome-wide significance for complex traits, as a function of sample size**

Lancet Neurol 2014;13:405-18