Familial Breast Cancer
SEARCHING THE GENES

Samuel J. Haryono

Results
BRCA Mutations

Verified Germline BRCA1 and BRCA2 Mutations

No.	Sample No.	Age of Onset	FH	Gene	Exon	Nucleotides	Amino Acid	Mutation Type	Validation	BIC
1.	AF+/T/JO-111	40	+	BRCA1	15	c.4956A>G	p.Met1652Ile	Missense	p.Ser1613Gly	204
2.				BRCA2	11	c.3627>A	p.Leu824X	Nonsense	p.Leu824X	NC
3	AF+/T/JO-166	33	+	BRCA2	27	c.10461A>T	p.Ile3412Va	Missense	p.Ile3412Val	109
4.	AF+/T/JO-117	35	-	BRCA2	27	c.10461A>T	p.Ile3412Va	Missense	p.Ile3412Val	109
5.	AF+/T/JO-206	37	-	BRCA2	27	c.9975>T	p.Lys3326X	Nonsense	p.Lys3326X	NC
6.	AF+/T/JO-100	44	-	BRCA2	11	c.3672A>G	p.Met1149Val	Missense	p.Met1149Val	5
7.	AF+/T/JO-134	39	+	BRCA2	11	c.3672A>G	p.Met1149Val	Missense	p.Met1149Val	5
8.	AF+/B/JK-06	38	+	BRCA2	11	c.4600T>C	p.His1458Try	Missense	p.His1458Try	3
9.	AF+/B/JK-156	34	+	BRCA2	18	c.8293T>C	p.Leu2698Pro	Missense	p.Leu2698Pro	NC
10.	AF+/W/JK-52	58	+	BRCA2	17	c.8110T>G	p.Ile2627Ser	Missense	p.Ile2627Ser	NC

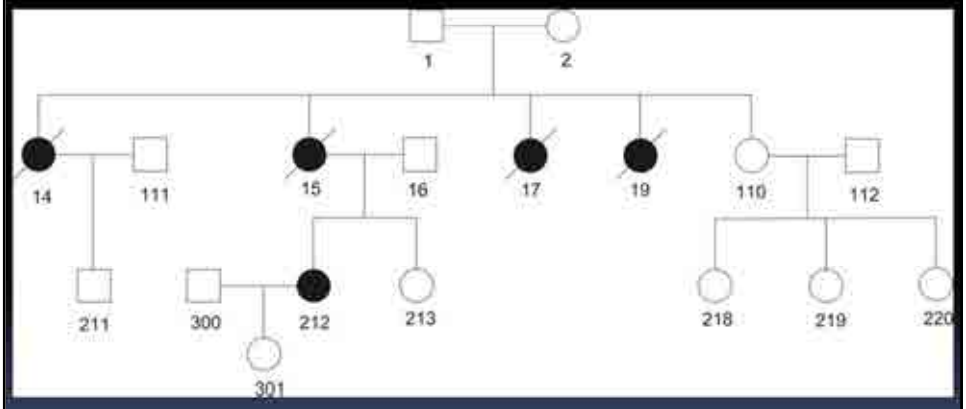
Issues in HSBOC

- Spectrum of mutation testing in familial breast cancer
- Variant of BRCA vs mutation of BRCA
- Clinical guideline and management for carrier
- BRCA mutation with TNBC
- Recurrent mutation in Indonesia
- Penetrance , sporadic and modifier genes
- Sanger, HRM, NGS
- Options of preventive, risk assessment, genetic counselling
- SAMANDA (Syndrome-Adopted Mutation-Assessment-Need-Data-Analysis) risk calculation for Indonesian specific FBC

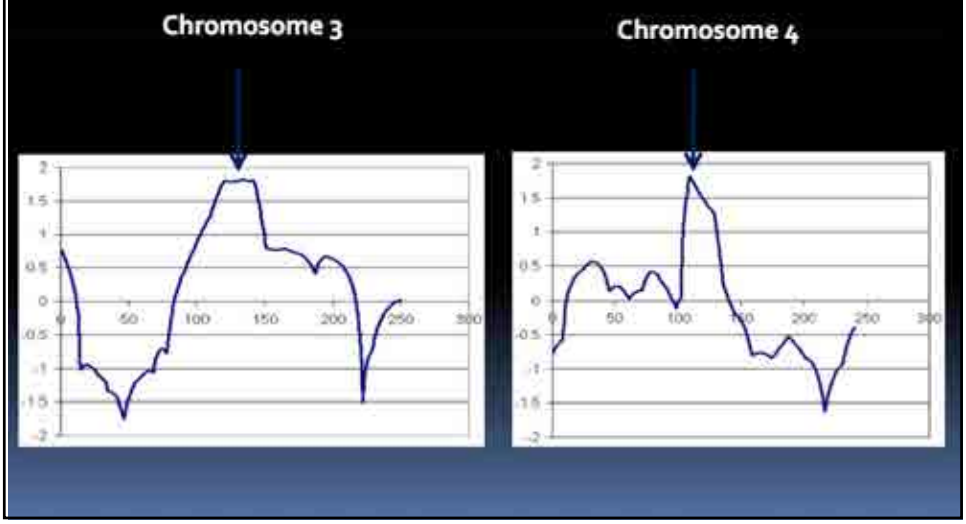
Methods: Linkage Analysis

- Genome-wide linkage analysis was done on 14 samples from one family (family code number A14) using 400 DNA repeat markers located spanning throughout genome.
- Analysis was done using ABI 3700 DNA sequencers, the result were then processed using Genemapper software. The assumed model used was autosomal dominant model.

Linkage Analysis – 400 STR markers on sample A14 family

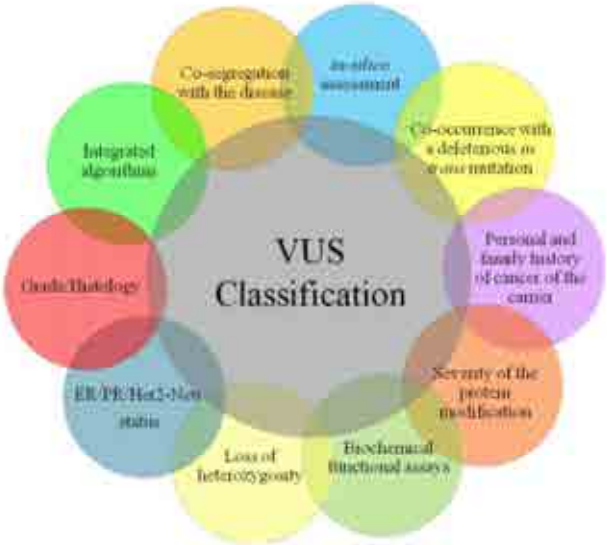


Linkage Analysis – 400 marker STR



Summary

- Several BRCA1/2 mutations in Indonesian familial breast cancer cases were found.
- There were 4 novel mutations found, all has been reported to the BIC
- We expected the presence of other predisposition genes
  - Confirmed with chromosome 3 and 4 showing fair evidence of linkage in linkage analysis.



Calo V, Bruno L, La Paglia L, et.al. 2010. The Clinical Significance of Unknown Sequence Variants in BRCA Genes. Cancers 2, 1644–1660

Components Assessed:

- LOH*
- Functional Domain
  - Unconserved
  - Conserved → In Silico Analysis (A-GVGD Class)
- Splice Site Analysis (ESE finder)
- Co-occurence in trans with other deleterious mutation
- Age at Diagnosis
- Grade
- ER - PR - HER-2 status
- Histopathologic Type

Odds Used for VUS Classification

No LOH	0.285	0.428
Loss of wild type	5.26	4.6
Loss of variant	0.067	0.067
Unconserved domain	0.01	0.01
Conserved C0	0.01	0.01
Conserved C15, C25	0.41	0.41
Conserved C35, C45, C55	1.5	1.5
Conserved C65	4.26	4.26
Splice defect	1,000	1,000
Not affecting splicing	0.01	0.01
Truncating mutation	1,000	1,000
In trans with mutation	0.0001	0.001

Odds Used for VUS Classification

Diagnosis at 50–59 years	1.67	2.07
Diagnosis at 40–49 years	3.40	2.89
Diagnosis at 30–39 years	9.65	4.97
Diagnosis at < 30 years	15.3	4.71
Grade 1	0.11	0.77
Grade 2	0.55	NA
Grade 3	1.97	1.3

Odds Used for VUS Classification

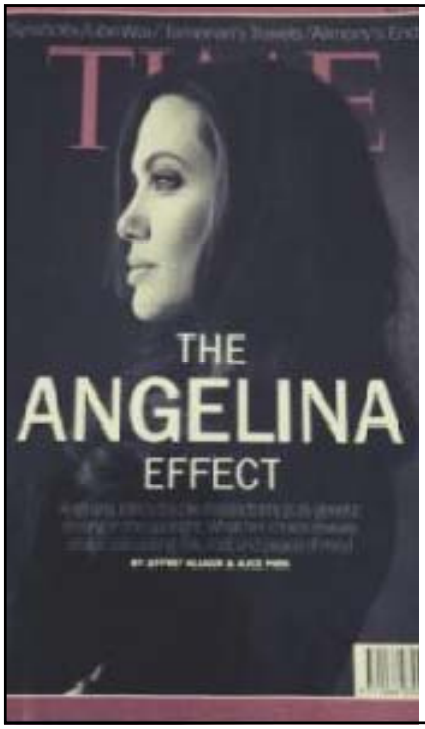
Medullary	8.0	NA
Triple negative	5.0	NA
ER positive	0.23	NA
ER negative	3.2	NA
PR positive	0.31	NA
PR negative	2.29	NA
HER-2 positive	0.15	0.15
HER-2 negative	1.2	1.2
ER positive/grade 1	0.067	0.23
ER positive/grade 2	0.37	1.5
ER positive/grade 3	NA	2.2
ER negative/grade 1	0.2	0.4
ER negative/grade 2	NA	0.38
ER negative/grade 3	4.1	NA

Sequence Change	LOH	Conserved Domain	A-DVGD Class	Splice Defect	In Trans	Age at Diagnosis (years)	Grade	ER Status	PR Status	HER-2 Status	Histology
BRCA1											
1294del40	WT	NA	NA	ND	NA	40	3	+	+	+	IDC
IVS13+1 G>A	None	NA	NA	Yes	No	41	3	+	+	–	IDC
C61G	WT	Yes	C65	ND	No	29	3	+	–	–	IDC
Y856H	None	No	C0	ND	Yes	46	3	+	+	+	DCIS
BRCA2											
2041delA	None	NA	NA	ND	NA	34	3	+	+	–	IDC
IVS8-12delTA	None	NA	NA	No	No	44	1	+	+	–	IDC
N817S	None	No	C0	ND	No	45	1	+	+	ND	IDC
L2109P	None	No	C0	ND	No	57	3	+	+	–	IDC

Sequence Change	Odds															
	LOH	A-DVGD/ Mutation	Splice	In Trans	Age at Diagnosis	Triple Negative	ER Status	PR Status	HER2	Grade/ER	Grade	HP	Odds	Odds, no LOH	Interpretation
BRCA1															
1294del40	0.28	1.000	1	1	3.4	1	1	2.29	0.16	4.1	1	1	2.8×10^8	4,789	D
IVS13+1 G>A	0.286	1	1.000	1	3.4	6	1	1	1	1	0.55	1	2,666	9,500	D*
C61G	0.28	4.28	1	1	19.3	6	1	1	1	1	1.37	1	3,371	642	DS*
Y856H	0.286	0.01	1	0.0001	3.4	1	0.23	0.31	0.16	1	1.97	1	2×10^{-6}	7×10^{-6}	N
BRCA2															
2041delA	0.428	1.000	1	1	4.87	1	1	1	1.3	2.2	1	1	5.618	1.3×10^4	D
IVS8-12delTA	0.428	1	0.01	1	2.89	1	1	1	1.3	0.23	1	1	0.003	0.009	N
N817S	0.428	0.01	1	1	2.89	1	1	1	1	0.23	1	1	0.003	0.007	N
L2109P	0.428	0.01	1	1	3.07	1	1	1	1.3	2.2	1	1	0.023	0.054	U

Summary

- VUS is a problem commonly encountered during the detection of BRCA1/2 germline mutation testing
- An established system to classify VUS is needed to guide clinician perform risk prediction, carrier testing, and reproductive decision making
- Multifactorial model developed by Spearman et.al (2008) is a clinically feasible model to classify VUS with relatively high Sn and Sp



- 1866 Paul Broca →pedigree in BC varian mutation of BRCA1
  RRM =risk reduction mastectomy

  in Asia is less: brca2, ER+, lower penetrans, less evidence,

  is it appropriate action / option?

GWAS

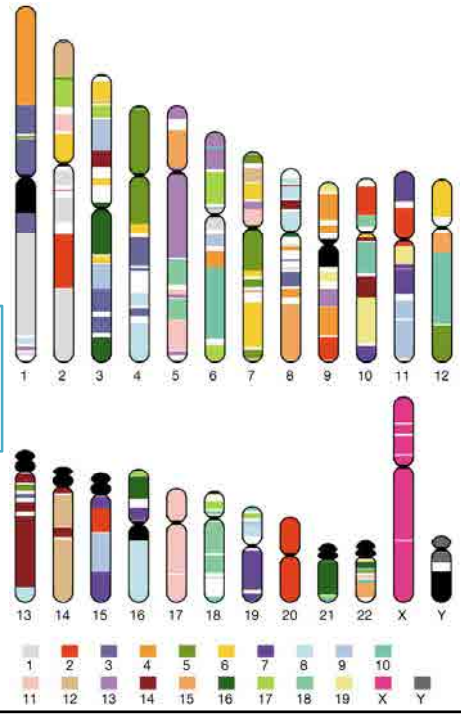
(Genome-Wide Association Study)

Introduction

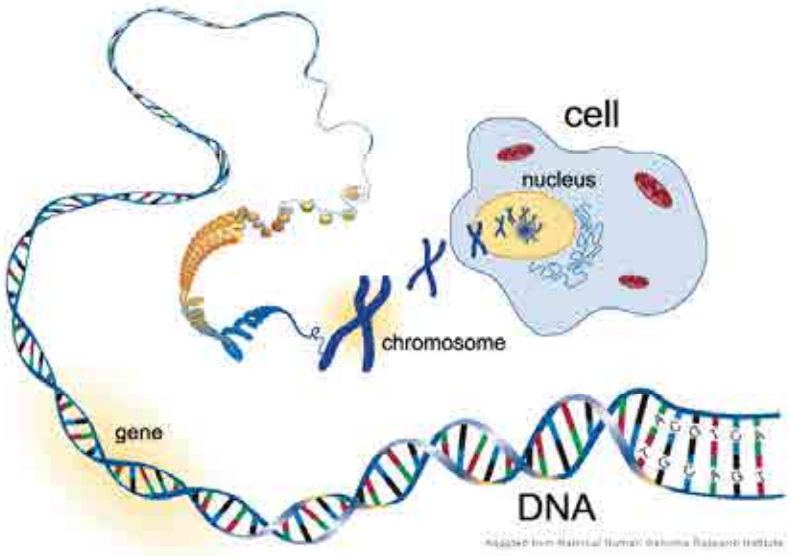
Development in genomic era

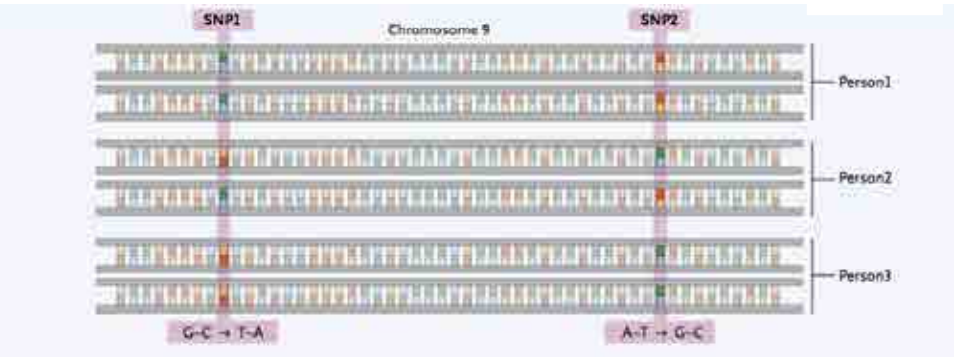
+

Not 100% Identical!!

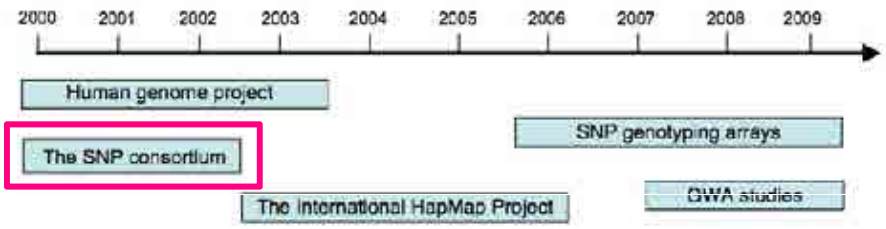


+



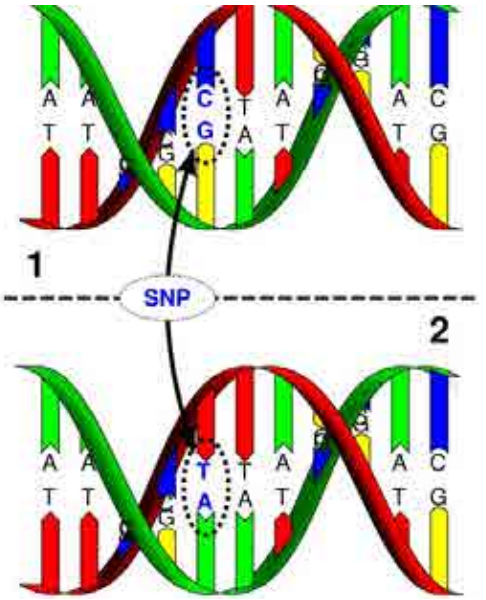


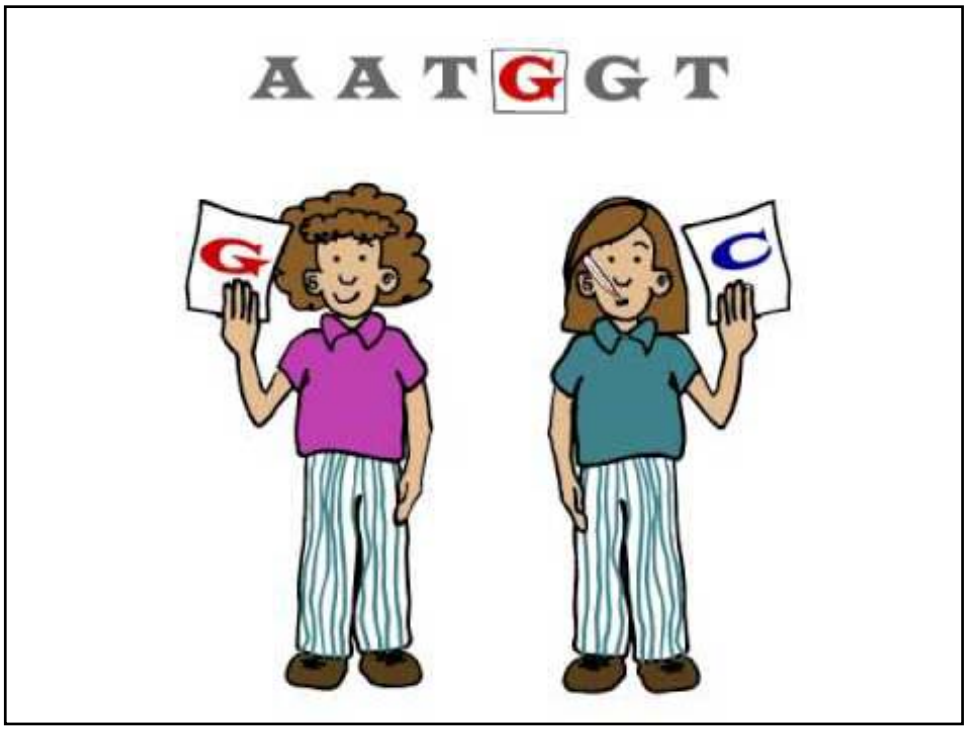
Development in genomic era



SNP

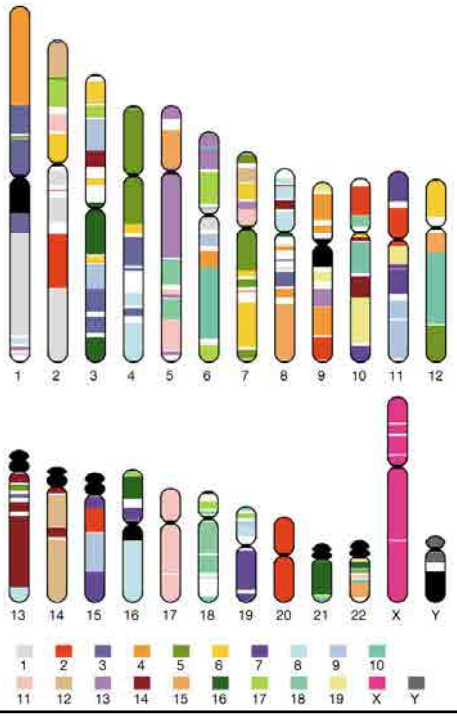
Single Nucleotide Polymorphism



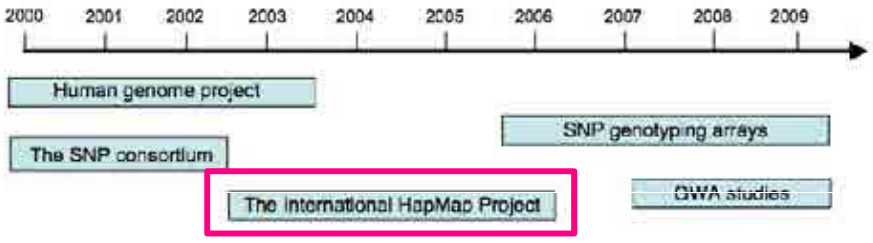


+

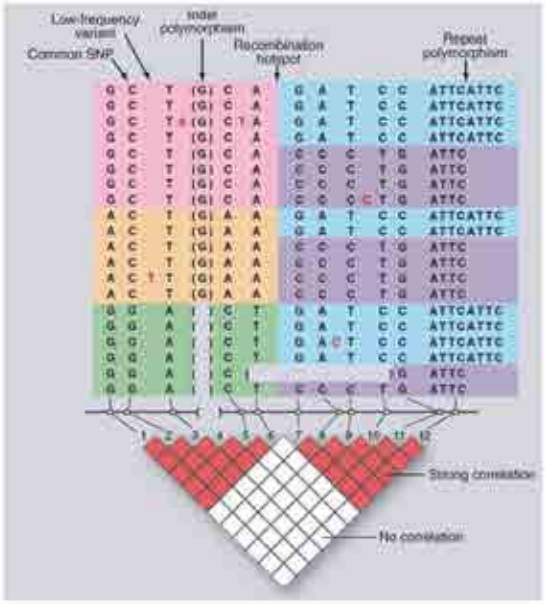
10 million SNPs



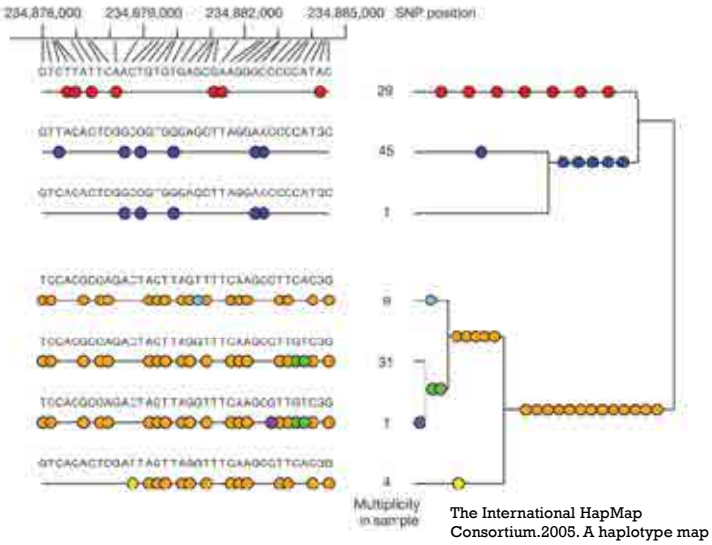
+ Development in genomic era



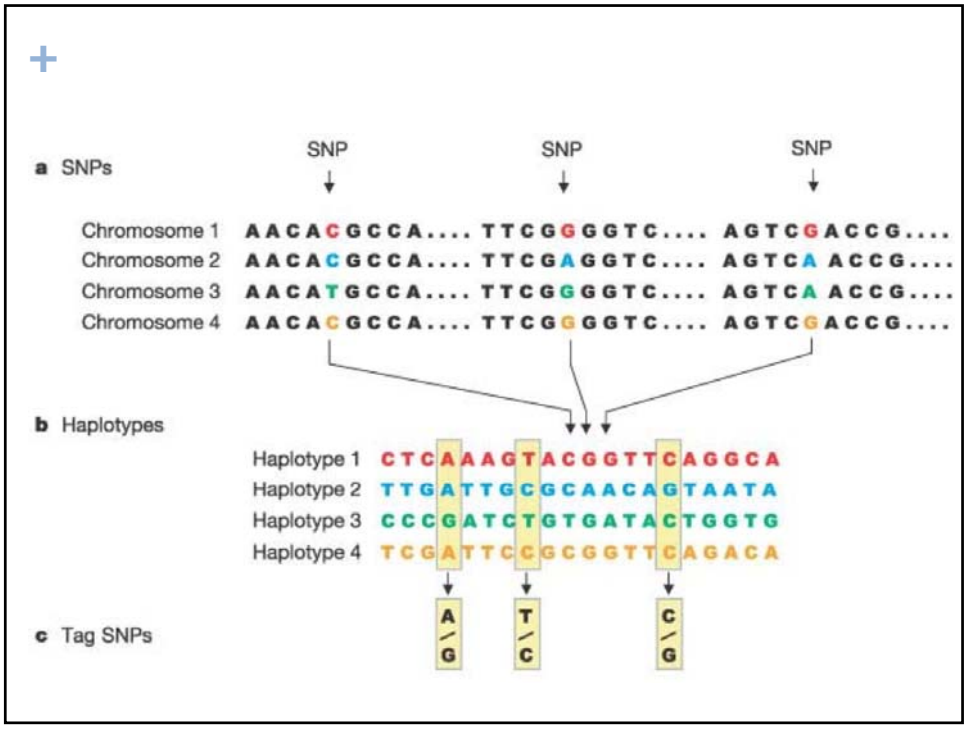
Lingkage SNPs



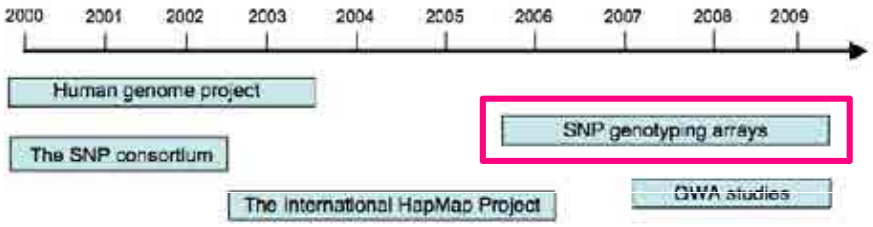
International Hapmap Consortium



The International HapMap Consortium.2005. A haplotype map of the human genome. *Nature*.437:1299-1320



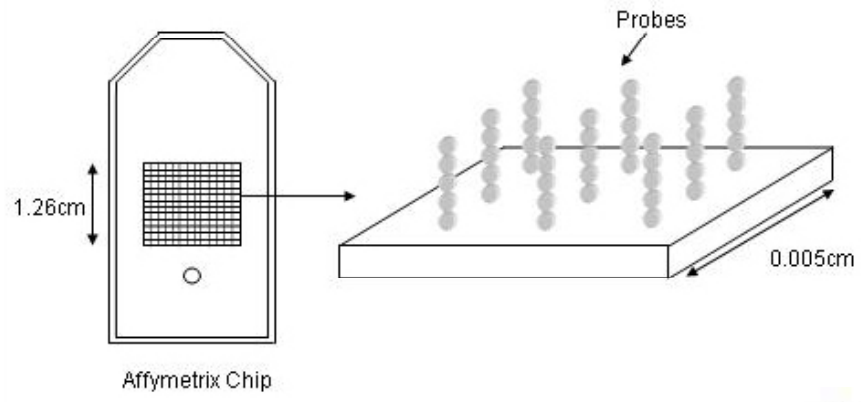
Development in genomic era

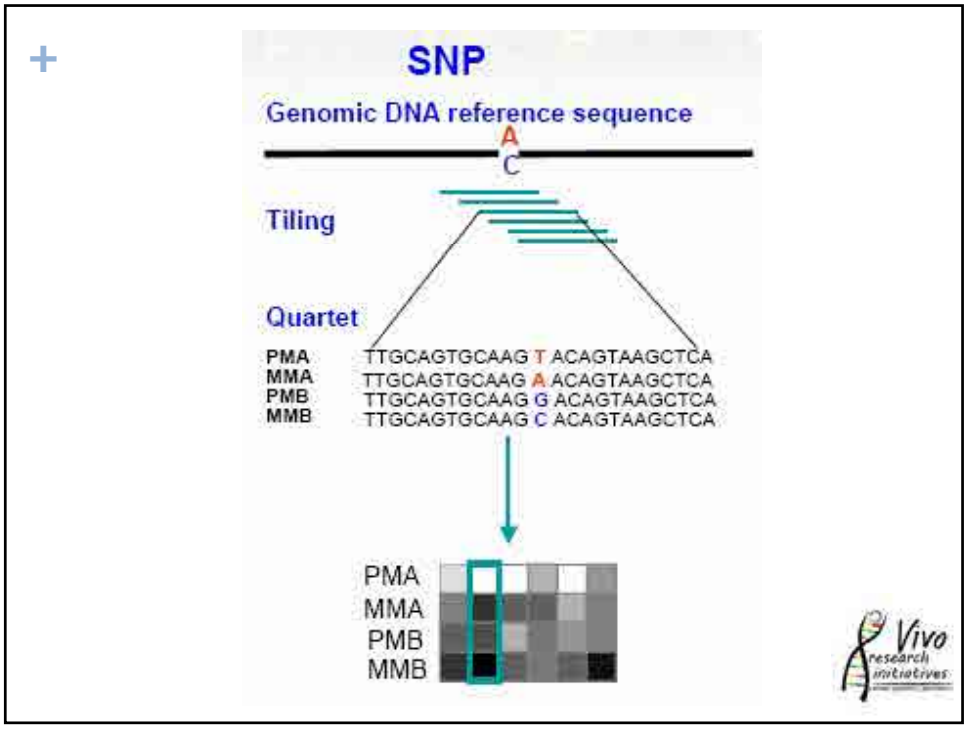


DNA Microarray

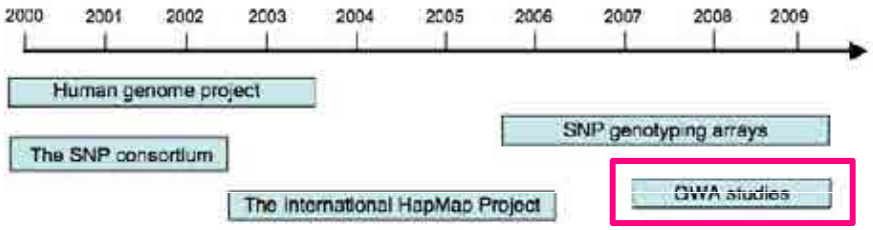
- Collection of microscopic DNA spots attached to a solid surface

Genome-Wide 5.0 Chip (Affymetrix)





Development in genomic era



GWAS
(Genome-Wide Association Study)

GWAS

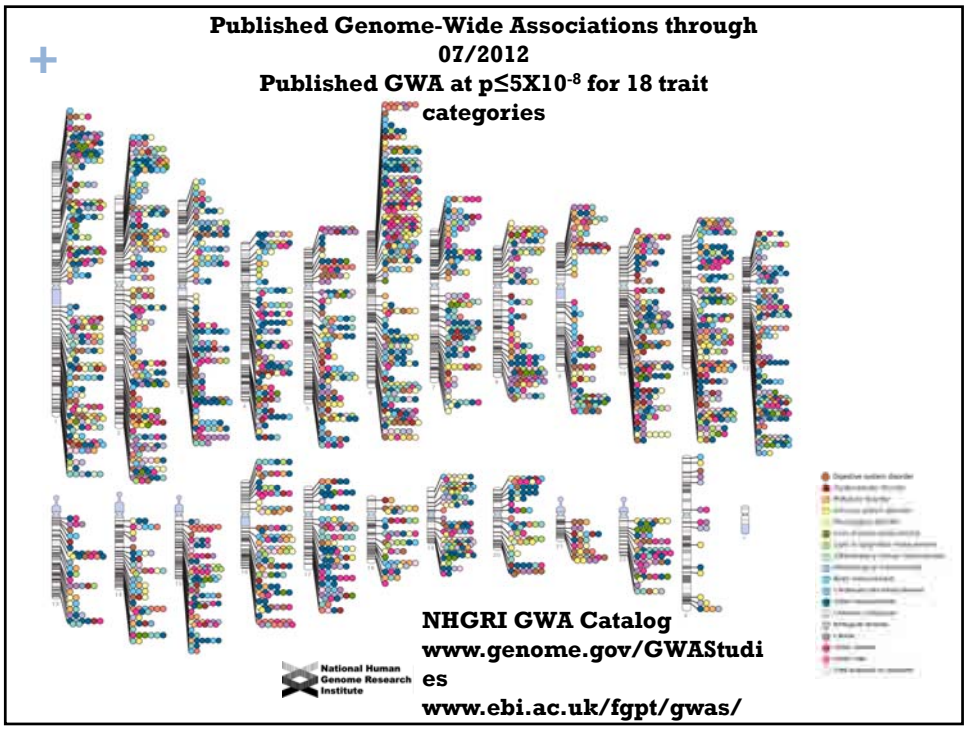
- Whole Genome Association Studies (WGAS)

- Correlation between **variant** vs **traits**

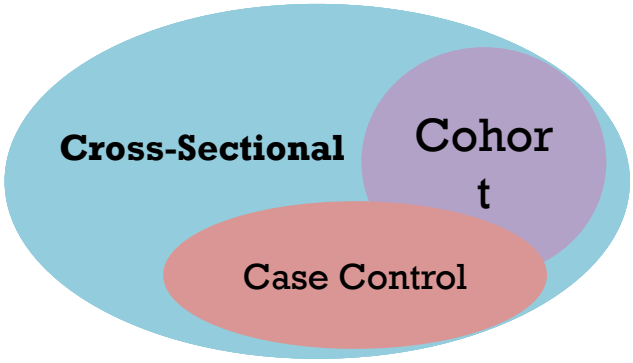
- **Variant** (e.g SNP) → genotype

- **Trait** (major disease in this case : breast cancer) → phenotype


Published Genome-Wide Associations through 12/2009.
658 published GWA at p≤5x10-8
NHGRI GWA Catalog
www.genome.gov/GWAStudies

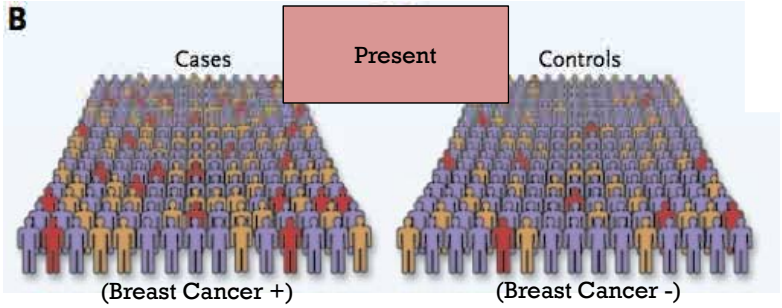

Published Genome-Wide Associations through 07/2012
Published GWA at p≤5X10-8 for 18 trait categories
NHGRI GWA Catalog
www.genome.gov/GWAStudies
www.ebi.ac.uk/fgpt/gwas/
National Human Genome Research Institute

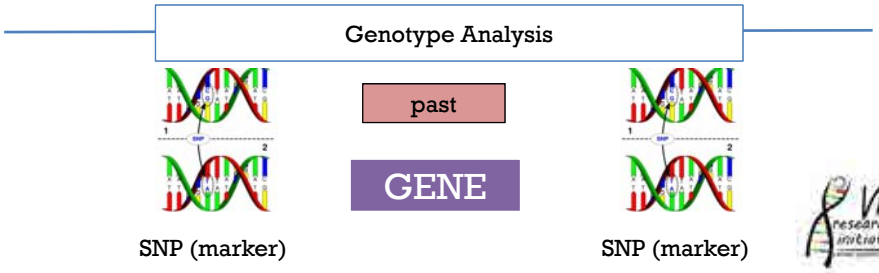
Association Study



	Case-control	Cohort	Cross-Sectional
Exposure	Past	Present	Presen
Disease	Present	Future	Presen

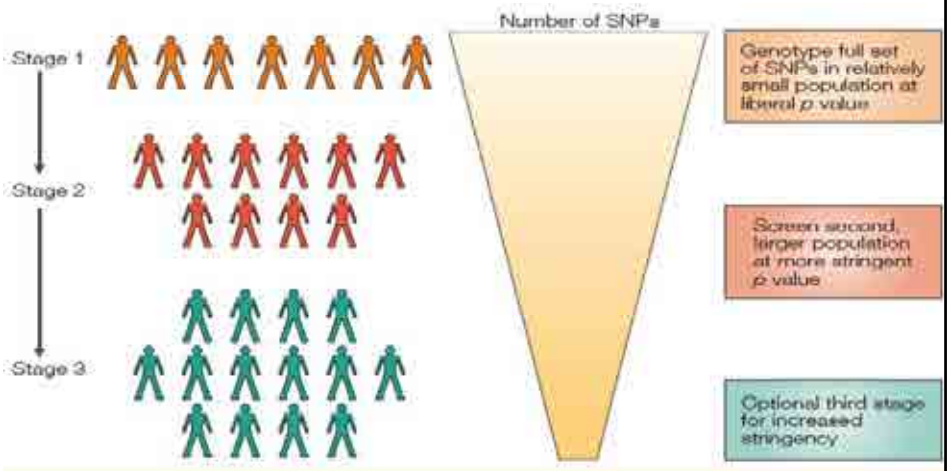
Case-Control

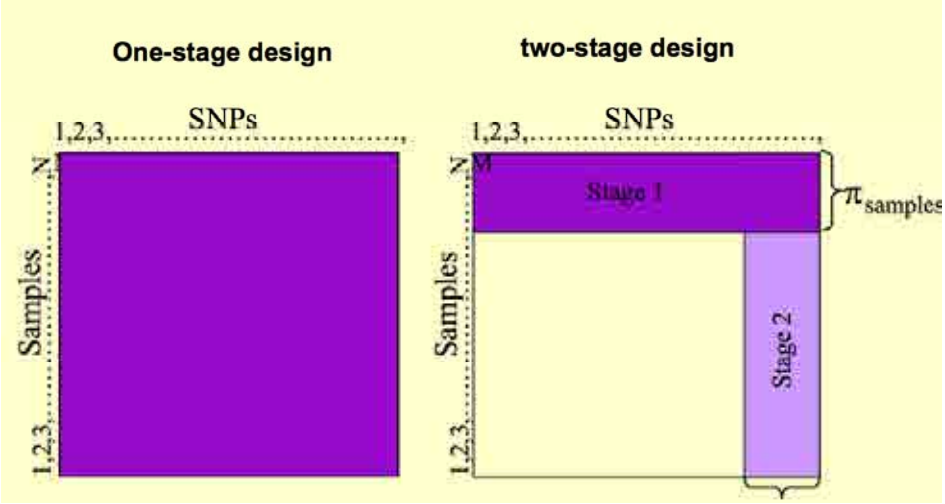




Methods

Multistage design





Sample Preparation

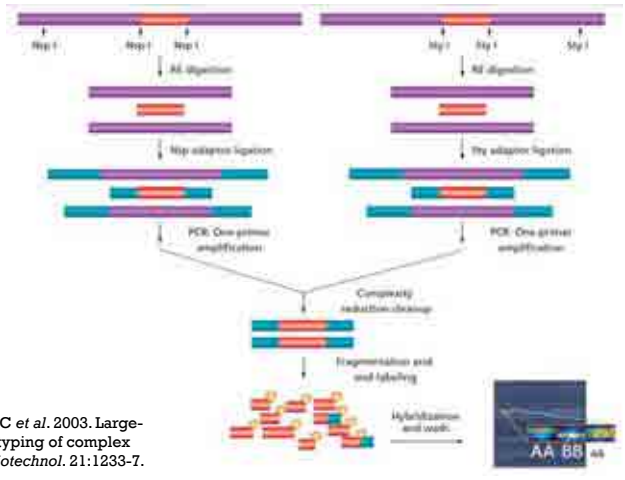
- 50 Cases
- 50 Control
- DNA Isolation (Roche)
- DNA quantity check ([DNA]= 100-200 ng/**µl**
- DNA quantity check (A260/A280 = 1.7-1.8)

Microarray

- Affymetrix©
- Genome-Wide SNP array 5.0

AFFYMETRIX®

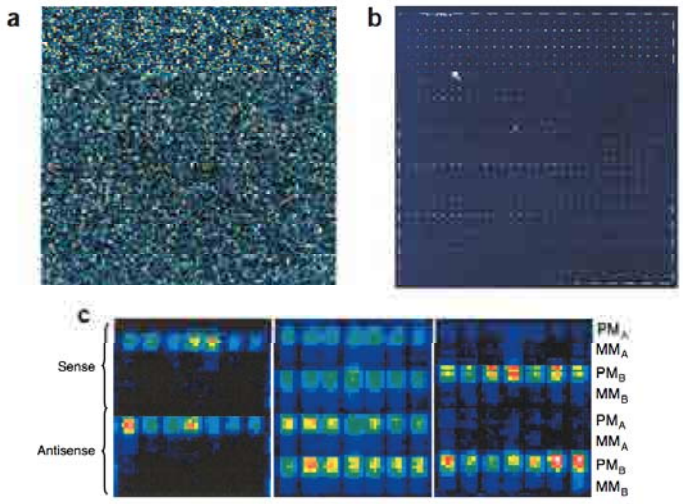
Sample Preparation (Whole Genome Sampling Assay)



Kennedy GC *et al*. 2003. Large-scale genotyping of complex DNA. *Nat Biotechnol*. 21:1233-7.

Vivo research initiatives

Signal Measurement

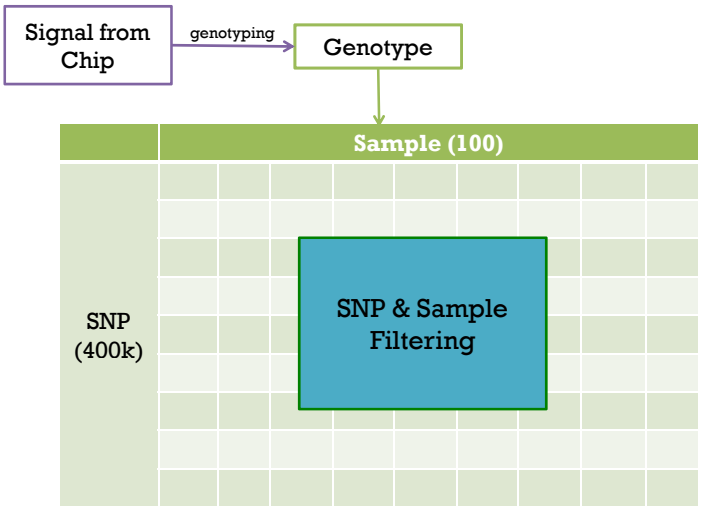


Kennedy GC *et al*. 2003. Large-scale genotyping of complex DNA. *Nat Biotechnol*. 21:1233-7.

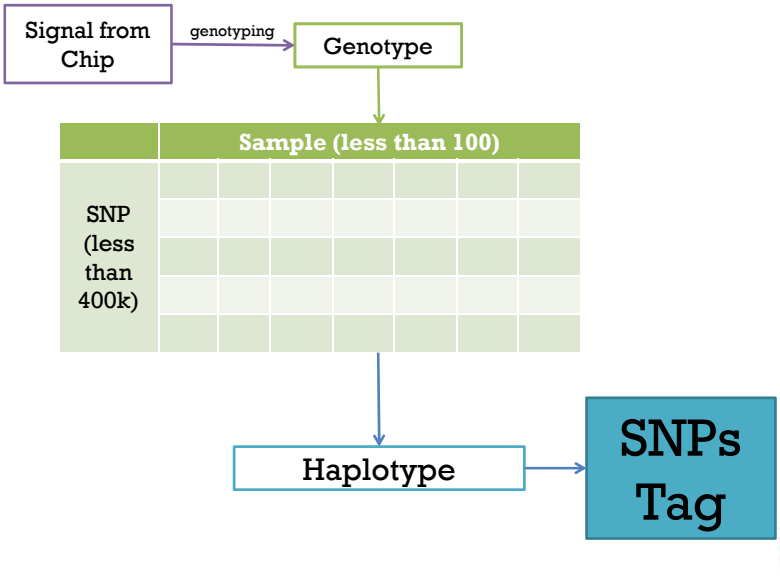
Data Analysis

Preliminary Analysis

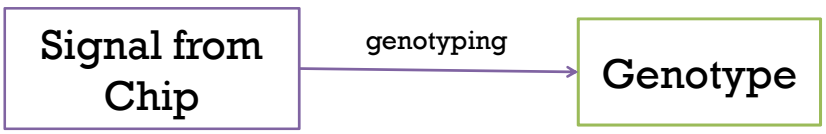
Flowchart



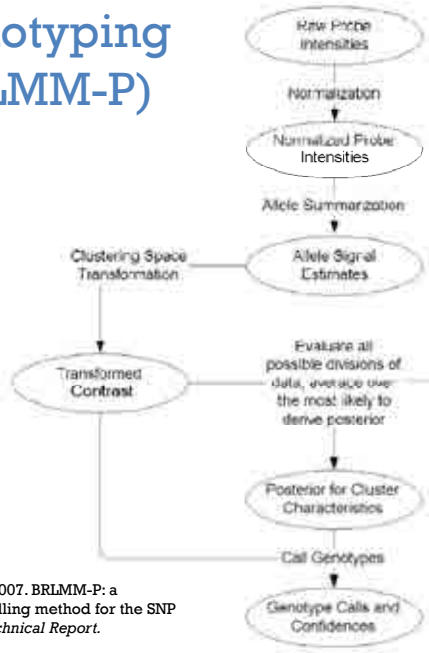
Flowchart



Genotyping



Genotyping (BRLMM-P)



Software : Genotyping Console 4.0 (Affymetrix)

Affymetrix.2007. BRLMM-P: a genotype calling method for the SNP 5.0 array. *Technical Report.*

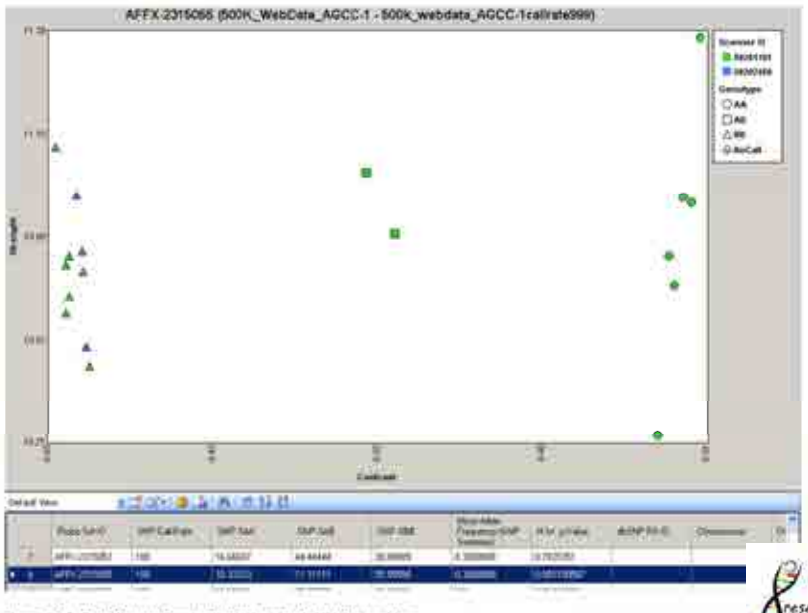
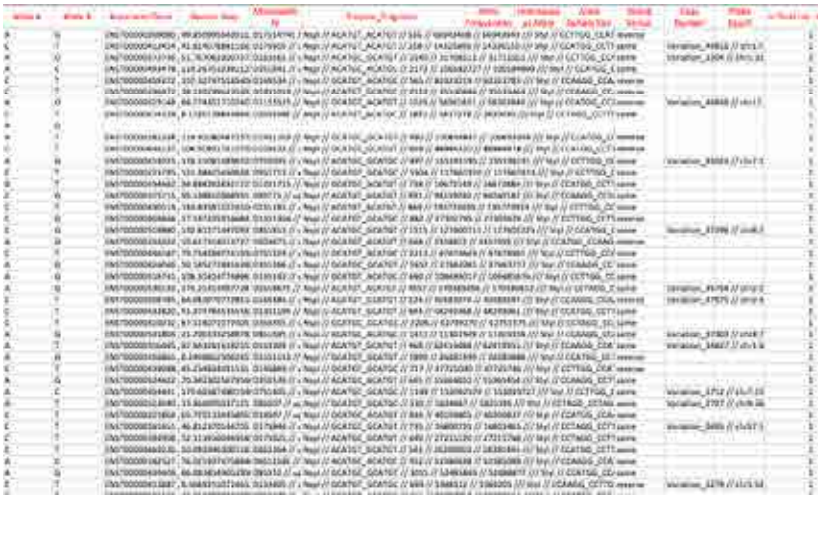


Figure 7.28. SNP Cluster Graph, BRLMM and BRLMM-P algorithm

+ SNP summary (1)

Index	Probe Set ID	SNP Call Rate	SNP %AA	SNP %AB	SNP %BB	Minor Allele Frequency	H.W. p Value	dbSNP RS ID	Chromosome	Chromosome Start	Chromosome Stop	Strand	ChrX pseudo	Cytoband	Flank	Allele A	Allele B
1	AFFX-SNP_1000	100	35.95506	50.5618	13.48315	0.3876404	0.585697	rs414795	17	66943738	66943738	+		q24.2	GGATAAGGATGGGCTA[A/G]ATTATCATTGCTGTTA	A	G
2	AFFX-SNP_1000	100	42.69663	51.68539	5.617978	0.3146068	0.008096	rs202221	17	14326088	14326088	+		p32	GGATGAAACTAGAGGA[C/T]GTTATGGAAAATGAAA	C	T
3	AFFX-SNP_1001	100	42.13483	50	7.865169	0.3286517	0.075839	rs642572	1	31709555	31709555	+		p35.2	GGATGTTGGGAGAAACCA[A/G]GAATGCCACCCAGGTC	A	G
4	AFFX-SNP_1002	100	87.64045	12.35955	0	0.06179775	0.379514	rs1299571	2	106584554	106584554	+		q12.2	GGATTGATCATTATTA[A/C]CCCACTTTCAAGGTCA	A	C
5	AFFX-SNP_1002	100	2.247191	23.03371	74.7191	0.1376404	0.691777	rs9831841	16	82323721	82323721	-		q23.3	GGCAAAGTGAGGAAAA[C/T]GAACAGGGAATGAAGC	C	T
6	AFFX-SNP_1002	100	24.7191	47.75281	27.52809	0.4859551	0.555483	rs714248	14	95532970	95532970	-		q13.2	GGCAAGAGCCATGTAA[C/T]GTGGCCCAAAAGAAAA	C	T
7	AFFX-SNP_1003	100	14.04494	47.19101	38.76405	0.3764045	0.944222	rs194321	11	58362418	58362418	-		q12.1	GGCACGAAATAATAGC[A/G]CTCTCCTTCAACTTAT	A	G
8	AFFX-SNP_1003	100	0	5.617978	94.38202	0.02808989	0.699795	rs7271810	20	2628108	2628108	+		p13	GGCACTGAAACTGAAA[C/T]AGCACATTCCACTTCC	C	T
9	AFFX-SNP_1003	100	1.123595	26.96629	71.91011	0.1460674	0.28003								GGCACTGGCAATTTAA[A/G]TATTAGTTGTTTGCAG	A	G
10	AFFX-SNP_1003	100	5.05618	31.46067	63.48315	0.2078652	0.551268	rs801756	14	100655022	100655022			q32.2	GGCAGACATTAAGGGT[A/T]AAAGCTGGAGAATGAT	A	T
11	AFFX-SNP_1003	100	0	15.73034	84.26966	0.07865169	0.254736	rs127771	10	86964444	86964444	-		q23.1	GGCAGAGCACAGTAGA[C/T]ACAAGAAAAGCAGCCA	C	T
12	AFFX-SNP_1004	100	18.53933	48.8764	32.58427	0.4297753	0.970189	rs471654	7	155195631	155195631	+		q36.3	GGCAGGTTATGTGCAT[A/G]ATAGAAGGAAGAGCGG	A	G
13	AFFX-SNP_1004	100	10.11236	38.20225	51.68539	0.2921348	0.308613	rs328197	9	117665435	117665435	-		q32	GGCAGTCCTCCTTCTC[C/T]AGGTTACATGAGGCAG	C	T
14	AFFX-SNP_1004	100	33.70787	45.50562	20.78652	0.4353932	0.325672	rs107725	12	16672401	16672401	+		p12.3	GGCAGTGGATCCAAGA[G/T]AGCAGTGGGAACTGTA	G	T
15	AFFX-SNP_1004	100	34.26966	55.05618	10.67416	0.3820225	0.03674	rs703230	9	94259474	94259474	+		q22.31	GGCATAGTAACATGAA[C/G]GACAAAAGTCAGCCCA	C	G
16	AFFX-SNP_1005	100	2.247191	25.2809	72.47191	0.1488764	0.974152	rs980924	1	195774694	195774694	-		q31.3	GGCATGGTAATAGGCC[C/T]GAAAGGAGTCTTCTGG	C	T
17	AFFX-SNP_1005	100	15.36854	45.50562	39.32584	0.3792135	0.655067	rs477103	13	27303351	27303351	-		q12.13	GGCCAATGGAACATAA[C/G]CAGAGATGTTTCTAG	C	G
18	AFFX-SNP_1006	100	27.52809	48.31461	24.1573	0.4831461	0.863523	rs109561	8	127601843	127601843	+		q24.21	GGCCCAGTATTCTTGC[C/G]AGGCCTTTACTACAGA	C	G
19	AFFX-SNP_1007	100	32.02247	47.19101	20.78652	0.4438202	0.556178	rs613809	20	3338977	3338977	-		p13	GGCCCGGCTCTAAACC[A/G]GGCTCTGAATTTGGGC	A	G
20	AFFX-SNP_1007	100	2.808989	43.25843	53.93258	0.244382	0.022186	rs471873	7	87678271	87678271	+		q11.22	GGCCTACTGCAATACAG[C/T]AGTGATAATTATATG	C	T
21	AFFX-SNP_1007	100	16.29214	50	33.70787	0.4129213	0.676444	rs391310	2	27663711	27663711	+		p24.1	GGCCTCAAGACATGCAA[A/G]GGCTCATGTGACAGGT	A	G
22	AFFX-SNP_1007	100	97.75281	2.247191	0	0.01123596	0.879495	rs118484	13	106485046	106485046	+		q33.2	GGCCTCTAGTTCCAGA[A/G]GAGACAAGCAACAGAT	A	G
23	AFFX-SNP_1008	100	27.52809	48.31461	24.1573	0.4831461	0.663523	rs643312	2	170485551	170485551	+		q31.1	GGCTAAAGGAAAGGTC[A/G]CAAATCAGCTGTTAAC	A	G
24	AFFX-SNP_1009	100	2.247191	20.78652	76.96629	0.1264045	0.4327	rs406222	2	40483504	40483504	+		p22.1	GGCTCACTGGCCAAGA[C/T]ATACTTTCATGAATTC	C	T
25	AFFX-SNP_1009	100	11.79775	39.32584	48.8764	0.3146068	0.239756	rs656140	13	48293041	48293041	+		q14.2	GGCTCAGAGAACTTGG[C/T]CAACTCCCGAGGGGTG	C	T
26	AFFX-SNP_1009	100	56.74157	33.70787	9.550562	0.2640449	0.076696	rs224084	3	42750324	42750324	-		p22.1	GGCTCCTAAGGATGA[C/T]CAAAACCAATGTCTCA	C	T
27	AFFX-SNP_1010	100	10.11236	40.44944	49.4382	0.3033708	0.566079	rs300342	8	11303011	11303011	+		p23.1	GGCTGGATGGCTAAAC[A/G]AGCAGGCTGTGACTCA	A	G
28	AFFX-SNP_1010	100	6.741573	36.51685	56.74157	0.25	0.726503	rs373710	1	62419727	62419727	+		p31.3	GGCTGTTAACTATCAA[A/T]GTGTCTCATAAATGGA	A	T
29	AFFX-SNP_1010	100	2.808989	26.4045	70.78651	0.1601124	0.807651	rs116341	15	26082540	26082540	-		q12	GGCTTATTTCACTTCA[A/G]GTCCCTCAGGTTCATC	A	G
30	AFFX-SNP_1010	100	32.02247	46.62922	21.34831	0.4466293	0.449625	rs800747	14	47725348	47725348	-		q21.3	GGCTTCAATAAAATAGC[C/T]CATCAGAGTGAGATAT	C	T
31	AFFX-SNP_1011	100	1.123595	20.22472	78.65169	0.1123596	0.852625	rs107477	12	55364970	55364970	+		q13.2	GGCTTTGGAAAGAAAA[A/G]TGGAAACTATATGGTC	A	G
32	AFFX-SNP_1011	100	17.97753	43.82022	38.20225	0.2808764	0.250022	rs109524	7	153092662	153092662	+		q36.2	GGGAAAAG[T/G]CACCCA[A/C]ATCTTGTCTACACAAT	A	C
33	AFFX-SNP_1011	100	15.73034	38.20225	46.06742	0.3483146	0.034449	rs270075	8	5825194	5825194	-		p23.2	GGGAAACAATCAAATA[G/T]GACAGGATTTTCACAGA	G	T
34	AFFX-SNP_1011	100	10.67416	47.19101	42.13483	0.3426966	0.526267	rs104092	19	40200508	40200508	+		q13.2	GGGAAATATTGGAGCC[C/T]AGTGTTTTCCTTAAAG	C	T
35	AFFX-SNP_1011	100	6.741573	32.58427	60.67416	0.2303371	0.279818	rs721403	17	16800880	16800880	+		p11.2	GGGAAGCCTGGAACGA[C/T]GTAATGGGCAAGGAGG	C	T
36	AFFX-SNP_1012	100	54.49438	41.57303	3.932584	0.247191	0.11844	rs721640	17	27215702	27215702	+		q11.2	GGGACAGAAGTAATGC[C/T]ACCTCCAGTGGAAAAA	C	T
37	AFFX-SNP_1012	100	1.123595	11.23596	87.64045	0.06741573	0.155638	rs375714	6	26381316	26381316	+		p22.2	GGGACCTACTGGTTCA[C/T]CCTTGCTGGTCCTTGA	C	T
38	AFFX-SNP_1013	100	43.25843	46.06742	10.67416	0.3370787	0.681217	rs471206	6	52385064	52385064	+		p12.2	GGGAGATCAGGAAGCA[A/C]CCCTTAGAGGTCAGGG	A	C
39	AFFX-SNP_1013	100	11.23596	44.94382	43.82022	0.3370787	0.939915	rs227248	12	52485891	52485891	+		q13.13	GGGAGGGCGTTCCAAA[A/G]GCTATGAAAACATTGG	A	G
40	AFFX-SNP_1015	100	52.24719	37.07865	10.67416	0.2921348	0.167404	rs943952	1	5368769	5368769			p36.32	GGGCAGTGGGAAATTC[C/T]ATGGTTCCA		

+ SNP summary (2)

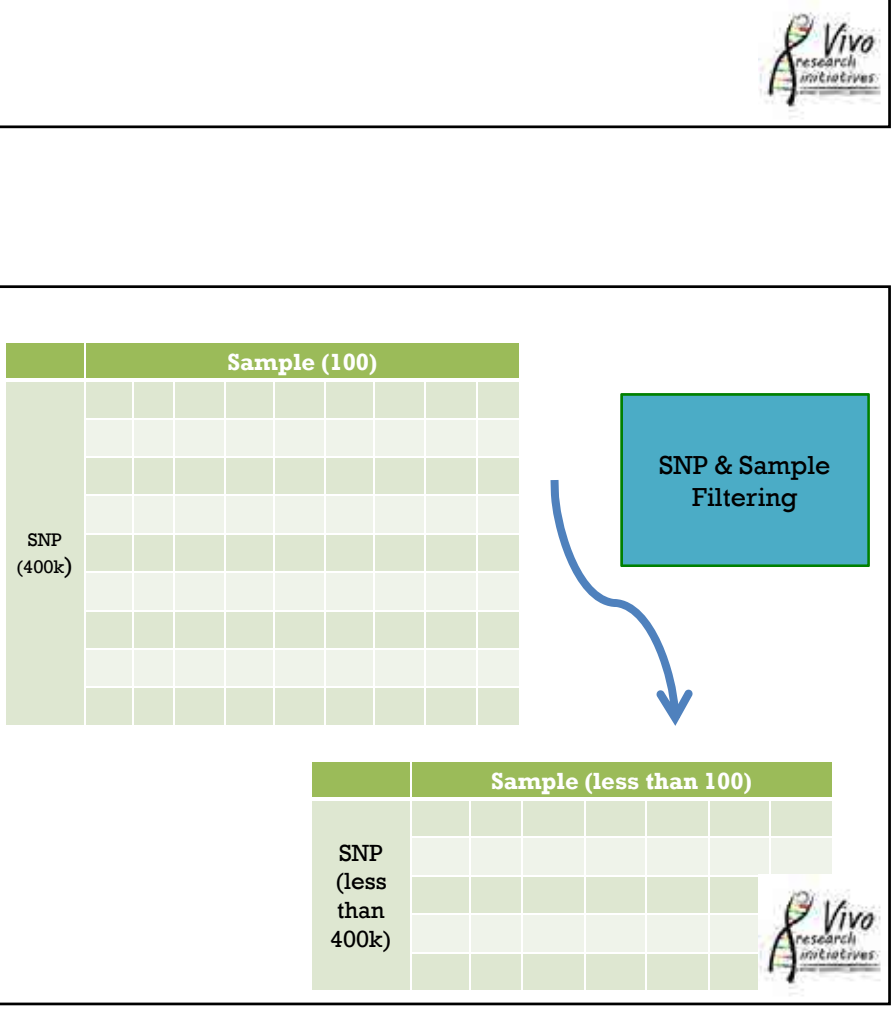


+ List of SNPs per-sample and its genotype

Probe Set ID	Call Codes	Confidence	Forced Call Codes	Contrast	Strength	Forward Strand Base Calls	Forced Forward
Strand Base Calls	dbSNP RS ID	Chromosomal Position	Affy SNP ID				
SNP_A-2058485	BB	0.000	BB	-0.670	9.003	GG	GG	rs2378338	120878622
SNP_A-2313216	AB	0.000	AB	-0.144	0.016	CT	CT	rs12870076	36672665
SNP_A-4196001	BB	0.000	BB	-0.601	10.272	TT	TT	rs1202778	58194231
SNP_A-1872630	BB	0.000	BB	-0.337	11.142	CC	CC	rs1202774	58485316
SNP_A-2380777	AA	0.000	AA	0.806	10.153	GG	GG	rs1202773	58404625
SNP_A-2097349	BB	0.000	BB	-0.657	10.602	AA	AA	rs2387831	75811769
SNP_A-2238870	BB	0.000	BB	-0.656	9.336	AA	AA	rs1796964	70215829
SNP_A-2145348	AA	0.000	AA	0.824	9.757	GG	GG	rs12758175	92678057
SNP_A-4250170	BB	0.000	BB	-0.793	10.956	AA	AA	rs1202761	90301592
SNP_A-1859386	AA	0.000	AA	0.606	9.804	AA	AA	rs34281 1201667
SNP_A-1948870	AA	0.000	AA	0.833	11.178	GG	GG	rs10852750	230982348
SNP_A-2096510	AA	0.000	AA	0.646	10.431	AA	AA	rs1249811	75863489
SNP_A-2047764	AA	0.000	AA	0.567	11.482	AA	AA	rs3189614	114523465
SNP_A-4285178	AA	0.000	AA	0.653	11.561	GG	GG	rs1252108	237165077
SNP_A-1983119	BB	0.000	BB	-0.677	10.115	TT	TT	rs10454054	212778587
SNP_A-1976384	AA	0.000	AA	0.817	10.371	CC	CC	rs1404871	75408868
SNP_A-2145276	AA	0.000	AA	0.784	9.143	GG	GG	rs1404232	75401683
SNP_A-2384157	AA	0.000	AA	0.797	10.473	GG	GG	rs11210478	75408411
SNP_A-2314251	BB	0.000	BB	-0.659	9.970	TT	TT	rs10838138	75404875
SNP_A-1936592	AA	0.000	AA	0.746	9.832	TT	TT	rs2551947	66948078
SNP_A-4248881	AB	0.000	AB	0.123	11.398	AG	AG	rs41352847	114189888
SNP_A-4234836	AA	0.000	AA	0.843	10.113	GG	GG	rs12062416	69925811
SNP_A-1983344	AB	0.000	AB	0.124	10.012	GA	GA	rs4767236	95703453
SNP_A-4278527	AB	0.000	AB	0.068	10.205	AG	AG	rs12041896	108253162
SNP_A-4203429	BB	0.000	BB	-0.388	10.003	AA	AA	rs475381	114062848
SNP_A-2125525	BB	0.000	BB	-0.699	9.862	TT	TT	rs41339545	16296695
SNP_A-4196549	AA	0.000	AA	0.689	9.152	CC	CC	rs4122277	86718870
SNP_A-4247941	AB	0.000	AB	0.223	9.589	CT	CT	rs6682847	214639253
SNP_A-2226171	BB	0.000	BB	-0.755	11.705	GG	GG	rs950894	16318788
SNP_A-2234286	BB	0.000	BB	-0.692	10.888	CC	CC	rs2574827	224945481
SNP_A-1986478	BB	0.000	BB	-0.684	12.037	GG	GG	rs3812162	224952159
SNP_A-4251286	BB	0.000	BB	-0.756	9.538	TT	TT	rs2802542	224958129
SNP_A-1964190	BB	0.000	BB	-0.668	10.637	AA	AA	rs2221956	119246934
SNP_A-2180686	AB	0.000	AB	-0.277	10.556	GA	GA	rs4699224	172654474
SNP_A-428248-	BB	0.000	BB	-0.674	10.988	AA	AA	rs13811490	106647411
SNP_A-4197886	AB	0.000	AB	0.181	10.624	GA	GA	rs968861	151761884
SNP_A-1988772	BB	0.000	BB	-0.753	11.038	GG	GG	rs1444541	231057343
SNP_A-2132690	AB	0.000	AB	0.151	9.794	AG	AG	rs868867	151761885
SNP_A-2139538	AB	0.000	AB	-0.119	11.144	AT	AT	rs11162392	322357392
SNP_A-1920000	AB	0.000	AB	-0.383	10.705	AG	AG	rs4512618	327252451

Probe set ID | Allele genotyping | confidence | Allele genotype | contrast | strength | Sequence genotyping | dbSNP rs ID | Chromosomal position

Sample & SNP Filtering



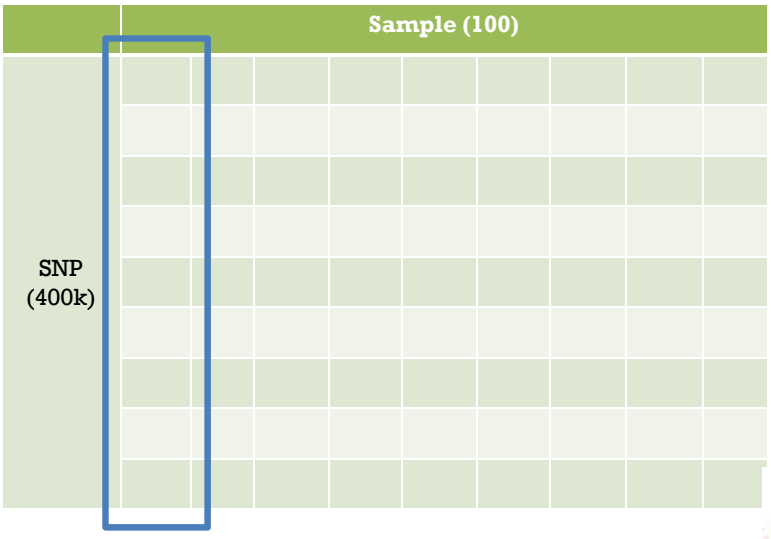
Sample Filtering

1. SNP Call Rate

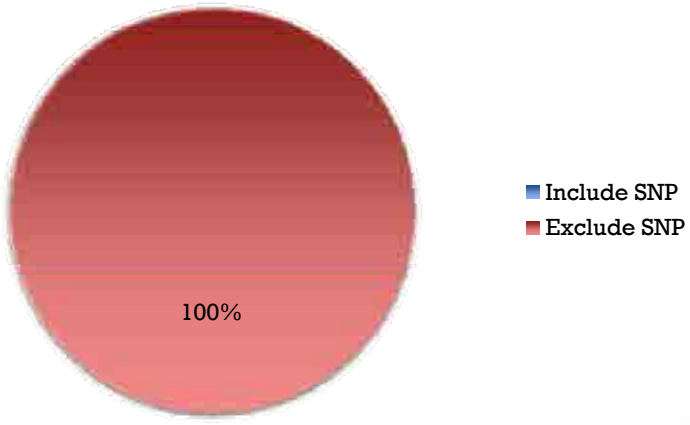
Fraction of called SNPs per sample over the total number of SNPs in the datasheet

Missing genotype data

Software :
Genotyping Console 4.0 (Affymetrix)



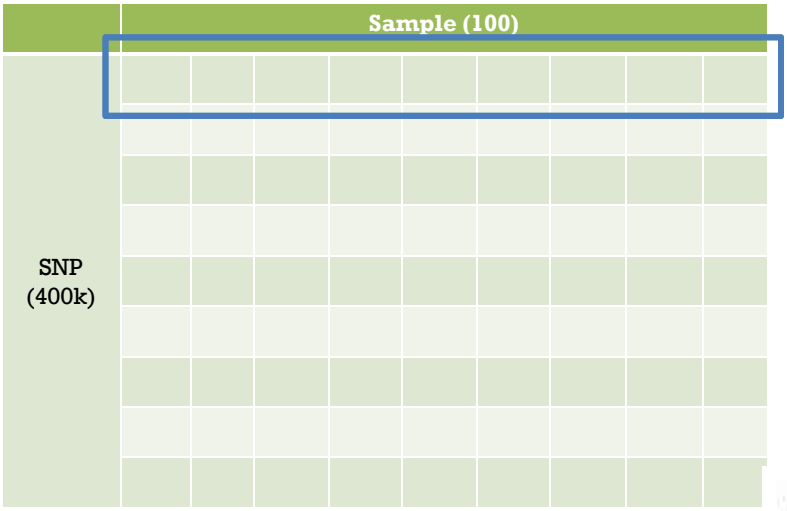
Sample Filtering



SNP Filtering

1. per-SNP Call Rate

Remove SNP per-SNP call rate (sometimes referred to as completeness) less than some threshold. Commonly-used value for the per-SNP call rate threshold range from 90-95%.



SNP Filtering

2. HWE (Hardy-Weinberg Equilibrium) p-value

Allele frequency vs Genotype frequency

Remove SNPs significantly out of Hardy Weinberg Equilibrium in cases and/or control. A p-value threshold in the range of 10^{-7} is commonly used.

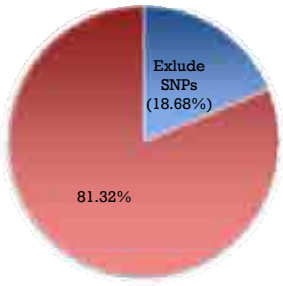
SNP Filtering

3. MAF (Minor Allele Frequency)

Frequency at which the less common allele occurs in a given population.

Conside removing SNPs with minor allele frequency (MAF) less than certain threshold (for example 1%)

Chromosom	SNPs	
	before	after
Chromosome 1	35402	28355
Chromosome 2	36786	29725
Chromosome 3	30222	24737
Chromosome 4	28628	22909
Chromosome 5	28572	23634
Chromosome 6	27954	23387
Chromosome 7	22816	18931
Chromosome 8	24498	19891
Chromosome 9	20139	16663
Chromosome 10	25170	20540
Chromosome 11	23092	18861
Chromosome 12	21960	18079
Chromosome 13	17054	13720
Chromosome 14	13950	11415
Chromosome 15	12703	10387
Chromosome 16	13464	10724
Chromosome 17	9872	7940
Chromosome 18	13236	10639
Chromosome 19	5475	4635
Chromosome 20	11043	8905
Chromosome 21	6268	5167
Chromosome 22	5334	4303
Chromosome X	9778	6974
total	**443416**	**360521**

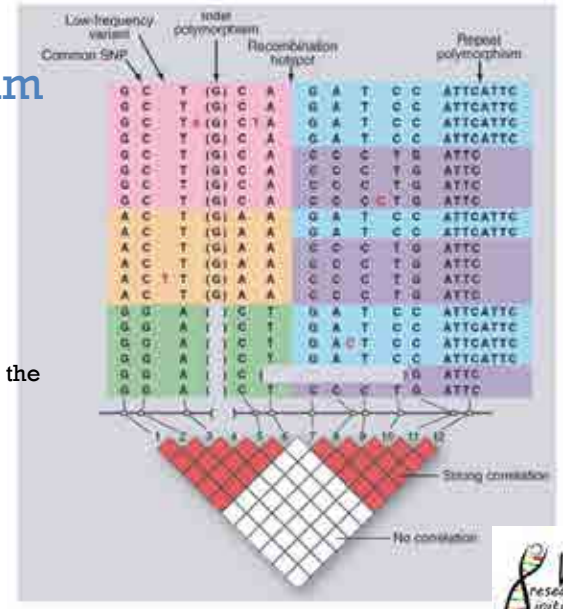


Haplotype Analysis

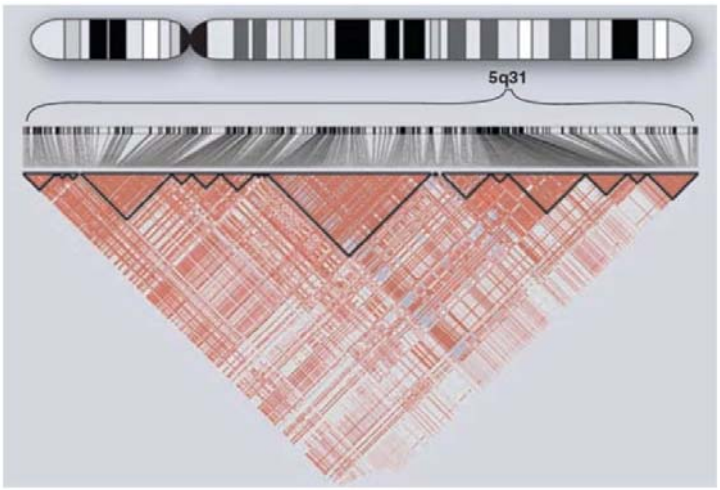
- Haplotype block
  - Pairwise tagging
  - Linkage Disequilibrium measurement→ r^2
- SNPs tags

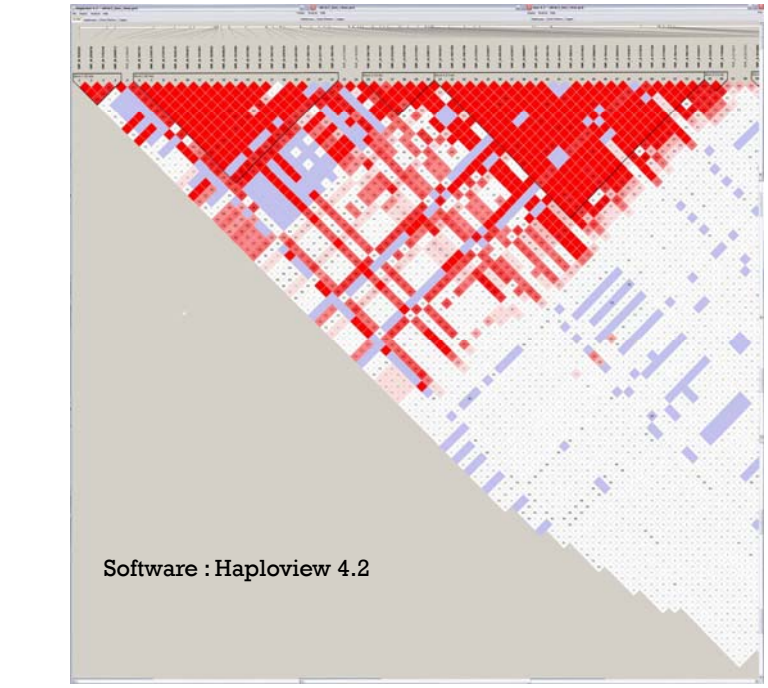
Linkage Disequilibrium

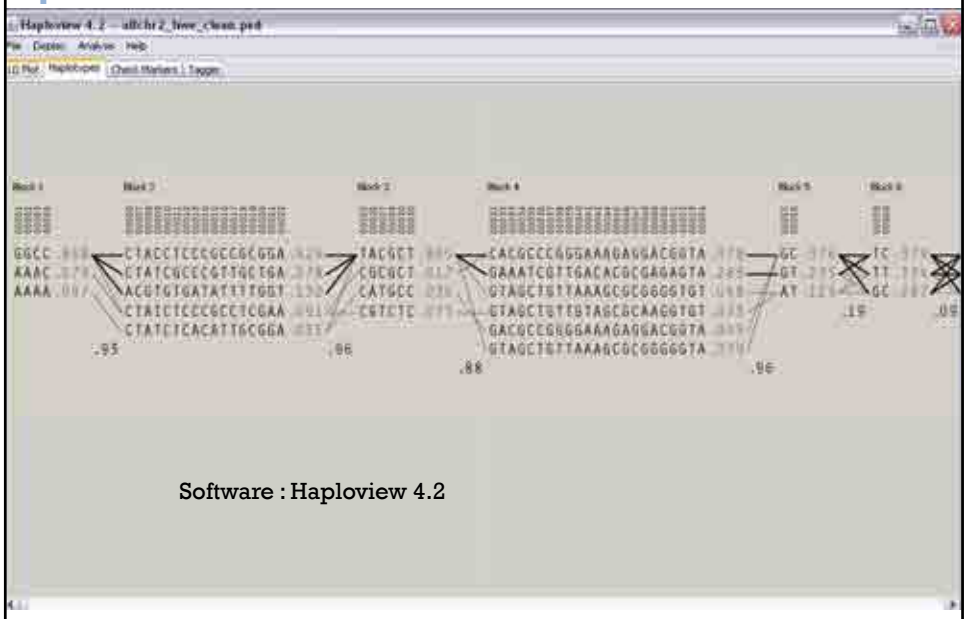
r^2 provides information on the corelation of two loci



Haplotype Block







Software : Haploview 4.2

Association Study

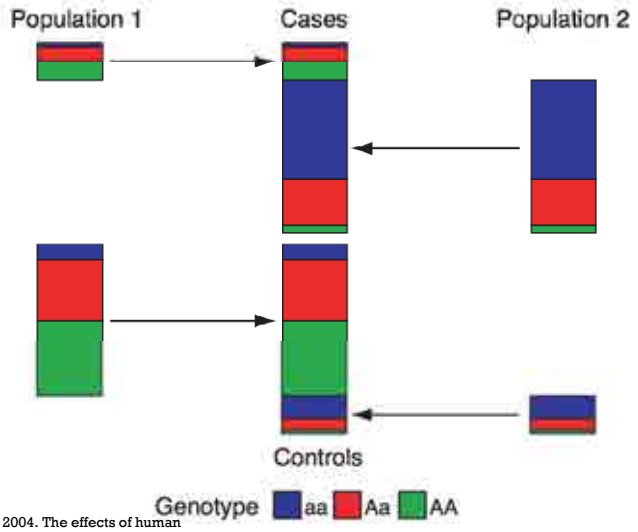
FOCUS ON STATISTICAL ANALYSIS

A tutorial on statistical methods for population association studies

David J. Balding

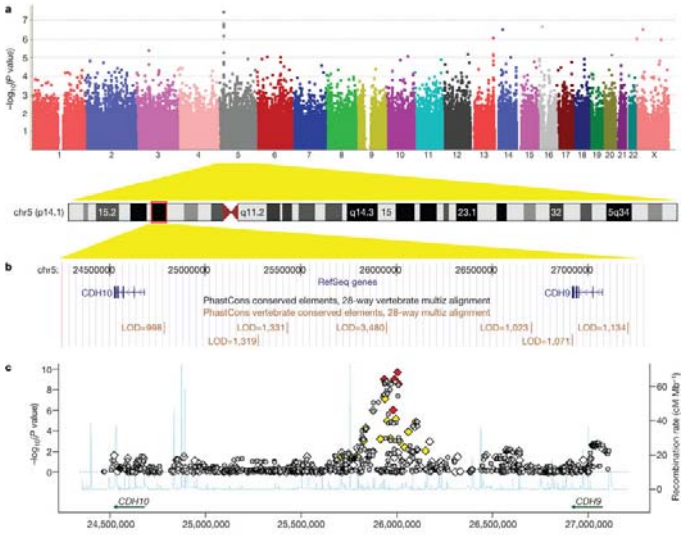
Abstract | Although genetic association studies have been with us for many years, even for the simplest analyses there is little consensus on the most appropriate statistical procedures. Here I give an overview of statistical approaches to population association studies, including preliminary analyses (Hardy–Weinberg equilibrium testing, inference of phase and missing data, and SNP tagging), and single-SNP and multipoint tests for association. My goal is to outline the key methods with a brief discussion of problems (population structure and multiple testing), avenues for solutions and some ongoing developments.

Population stratification



Marchini J, *et al*. 2004. The effects of human population structure on large genetic association studies. *Nat Genet*. 36:512-7

Manhattan Plot



Wang K, *et al*. 2009. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 459:528-533

Acknowledgement

Vivo research initiatives
cancer, genetics, genomics

Samuel Haryono
Althaf Setyawan
IGB Datasena
Wahyu Budi Santosa
Adrian Salim
Raymond Mulyarahardja
Dismas A. Chaspuri
Uti Nilam Sari
Kartika Sari

Thank You