Paolo Vineis

Imperial College London and ISI Foundation Torino

Field synopses of genetic evidence: DNA repair

A field synopsis on genetically-based low-penetrant defects in DNA repair and cancer

Paolo Vineis, Maurizio Manuguerra, Federica Saletta, Fotini K. Kavvoura, John P.A. Ioannidis, Giuseppe Matullo

We have created a regularly updated database of studies addressing associations between DNA repair gene variants (excluding highly penetrant mutations) and different types of cancer.

Website www.episat.org

Episat Molecular Epidemiology toolkit site - Windows Internet Explorer

http://www.episat.org/episat/index.php

File Modifica Visualizza Preferiti Strumenti ?

Episat Molecular Epidemiology toolkit site

ecnis episat project

Projects Lang login

- EpiSAT Home
- MEC database
- DNA Repair
- Reporting Standards and Databases for Omics Data
- About

EpiSat

EpiSAT is a web application which provides the instruments necessary to maintain and pool results of biomedical investigations. EpiSAT includes facilities for the storage, archiving, cataloguing, and disseminating pooled results. EpiSAT includes on-line analytical tools, powered by the R Statistical environment.

MEC - knowledge database

Construction and implementation of a knowledge data base on Molecular Epidemiology and Cancer. The projects currently included in MEC are shown in the windows below.

EpiSAT - MEC

Construction and implementation of a knowledge data base on Molecular Epidemiology and Cancer. The projects currently included in MEC are accessible following the link.

DNA Repair

In the DNA repair polymorphisms site an updated source of information on gene-environment interactions, and particularly on DNA repair and the risk of cancer, is available for geneticists and epidemiologists.

ecnis

National Institutes of Health NIEHS

EMBL-EBI

This site is maintained by ISI Foundation (F.Rosa, M.Manuguerra, F.Saletta, G.Matullo), University of New Mexico (M.Berwick) and CdC (M.Khoury) -

start Posta inviata - Ou... Episat Molecular E...

Episat Molecular Epidemiology toolkit site - Windows Internet Explorer

http://www.episat.org/episat/huge/

File Modifica Visualizza Preferiti Strumenti ?

Episat Molecular Epidemiology toolkit site

ecnis episat project

Projects Lang login

- EpiSAT Home
- MEC database
- DNA Repair
  - study design and comparative tables
  - all articles included
  - statistics
- Reporting Standards and Databases for Omics Data
- About

DNA Repair

In the DNA repair polymorphisms site an updated source of information on gene-environment interactions, and particularly on DNA repair and the risk of cancer, is available for geneticists and epidemiologists.

Studies included

Date	Study	Authors
2007	Genotypes, haplotypes and diplotypes of XPC and risk of bladder cancer *Carcinogenesis. 2007 Mar;28(3):698-703. Epub 2006 Oct 19.*	Zhu Y, Lai M, Yang H, Lin J, Huang M, Grossman HB, Dinney CP, Wu X.
2007	Association of the TP53 Codon 72 Polymorphism with Colorectal Cancer in a Chinese Population *Jpn J Clin Oncol. 2007 May;37(5):385-90*	Zhu ZZ, Wang AZ, Jia HR, Jin XX, He XL, Hou LF, Zhu G.
2007	A haplotype encompassing the variant allele of DNA repair gene polymorphism ERCC2/XPD Lys751Gln but not the variant allele of Asp312Asn is associated with risk of lung cancer in a northeastern Chinese population. *Cancer Genet Cytogenet. 2007 May;175(1):47-51*	Yin J, Vogel U, Ma Y, Qi R, Sun Z, Wang H.
2007	The DNA repair gene XRCC1 and genetic susceptibility of lung cancer in a northeastern Chinese population *Lung Cancer. 2007 May;56(2):153-60. Epub 2007 Feb 20*	Yin J, Vogel U, Ma Y, Qi R, Sun Z, Wang H.
2007	Association between polymorphisms of biotransformation and DNA-repair genes and risk of colorectal cancer in Taiwan *J Biomed Sci. 2007 Mar;14(2):183-93*	Yeh CC, Sung FC, Tang R, Chang-Chieh CR, Hsieh LL
2007	Genetic polymorphisms of the DNA repair gene and risk of nasopharyngeal carcinoma *DNA Cell Biol. 2007 Jul;26(7):491-6*	Yang ZH, Du B, Wei YS, Zhang JH, Zhou B, Liang WB, Jia J, Zhang BL, Zhang L
2007	Relationship between XRCC1 polymorphisms and susceptibility to prostate cancer in men from Han, Southern China. *Asian J Androl. 2007 May;9(3):331-8.*	Xu Z, Hua LX, Qian LX, Yang J, Wang XR, Zhang W, Wu HF
2007	Association between a functional single nucleotide polymorphism in the MDM2 gene and sporadic endometrial cancer risk	Walsh CS, Miller CW, Karlan BY, Koeffler HP

Internet

Episat Molecular Epidemiology toolkit site - Windows Internet Explorer

http://www.episat.org/episat/huge/?computation=2

File Modifica Visualizza Preferiti Strumenti ?

Episat Molecular Epidemiology toolkit site

ecnis episat project

Projects Langua

- EpiSAT Home
- MEC database
- DNA Repair
  - study design and comparative tables
  - all articles included
  - statistics
- Reporting Standards and Databases for Omics Data
- About

Do you want to submit a query to the NewHugeNet database?

Which gene:	All genes
Which polymorphism:	All
Which rs:	All
First author:	
Date:	-- Select a date --
Nationality:	All
Method:	-- Select an method --
Which gender:	(•) All ( ) Males ( ) Females
Cancer site:	-- Select a cancer site --
Which ethnic group:	-- Select an ethnic group --
Which histologic type:	-- Select an histologic type --
Study design:	-- Select an study design --

Submit

Internet

Episat Molecular Epidemiology tool indows Internet Explorer

http://www.episat.org/episat/huge/?com

File Modifica Visualizza Preferiti Strumenti ?

Episat Molecular Epidemiology toolkit site

ecnis episa

Projects

- EpiSAT Home
- MEC database
- DNA Repair
  - study design and comparative tables
  - all articles included
  - statistics
- Reporting Standards and Databases for Omics Data
- About

Do you want to su to the NewHugeNet database?

Which gene:	All genes, ALKBH3, APEX, ATM, ATR, BARD1, BLM, BRCA1, BRCA2, BRIP1, CCND1, CCNH, CD3EAP, CDKN1A, CHEK1, CHEK2, COMT, CYP19, DCLRE1A, DCLRE1B, DDB2, DMC1, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, ERCC6, ESR1, EXO1
Which polymorphism:	
Which rs:	
First author:	
Date:	--
Nationality:	All
Method:	-- Select an method --
Which gender:	All / Males / Females
Cancer site:	-- Select a cancer site --
Which ethnic group:	-- Select an ethnic group --
Which histologic type:	-- Select an histologic type --
Study design:	-- Select an study design --

Submit

Internet

ecnis episat project

Projects Lang ogin

- EpiSAT Home
- MEC database
- DNA Repair
  - study design and comparative tables
  - all articles included
  - statistics
- Reporting Standards and Databases for Omics Data
- About

Information currently available	
Investigators	295
Studies	417
Total subjects	449963
Controls	264441
Cases	185522

Distribution by cancer site

Site	Cases (n)	Controls (n)	Studies (n)
Head and neck	8584	8923	22
Lung	35496	90389	89
Bladder	9268	9489	22
Colorectal	16453	15814	28
Breast	61281	63323	83
Pancreas	693	1321	2
Cervix	7372	8401	49
Esophageal	3363	4678	17
Liver	1117	3129	5
Leukemia	3465	3967	9
Oral	1107	1424	7
Glioma	3556	3869	7
Skin	11167	11400	26
Pharynx	661	596	3
Larynx	338	350	2
Prostate	2680	2420	10
Ovarian	6337	10870	7
Lymphoma	3883	3598	6

Internet

Conjunctiva	125	125	1

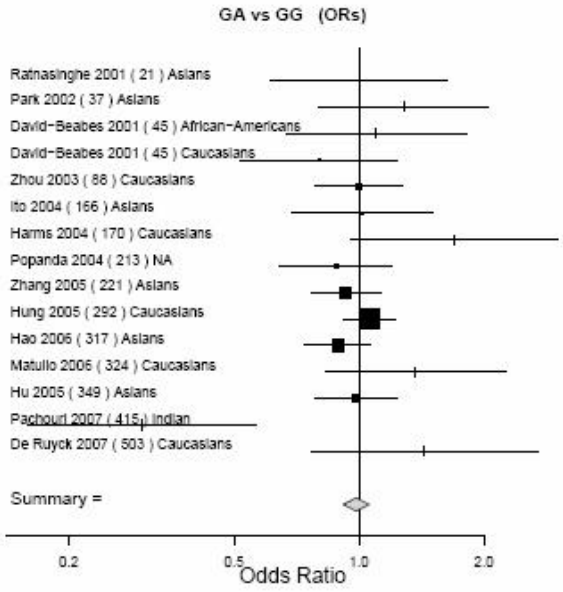
Distribution by genetic polymorphism

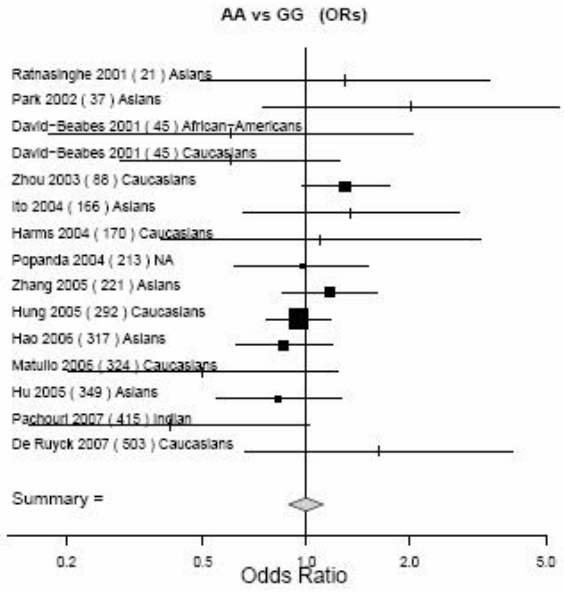
Gene	Cases (n)	Controls (n)	Studies (n)
XRCC1	50939	58145	132
XRCC2	21384	24496	18
XRCC3	35855	37685	73
ERCC2	39507	95941	99
ATM	14077	19292	11
BRCA1	13983	20384	13
BRCA2	20999	25278	18
CHEK2	11324	14395	8
XRCC6	3155	2834	3
LIG4	16380	16265	12
NBN	14134	16552	18
TP53	41772	49261	120
RAD51	8572	10785	13
RAD52	8680	11674	7
PARP1	12109	11213	14
ERCC1	13118	66567	30
OGG1	22288	21925	41
ERCC4	10417	11703	15
XPC	13960	15343	30
MGMT	11989	14126	22
LIG1	4486	4198	6
MLH1	7150	7599	8
XRCC5	1100	1137	3
XRCC4	11428	11308	7
ERCC5	16746	17323	24
XPA	9565	9599	15
PCNA	1439	1231	4

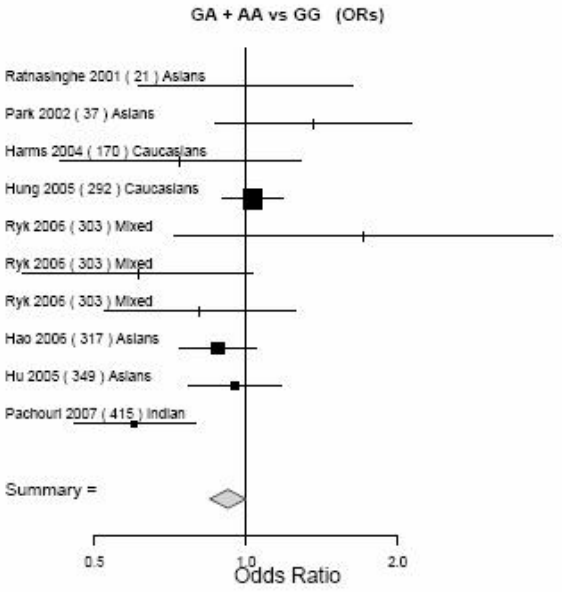
Internet

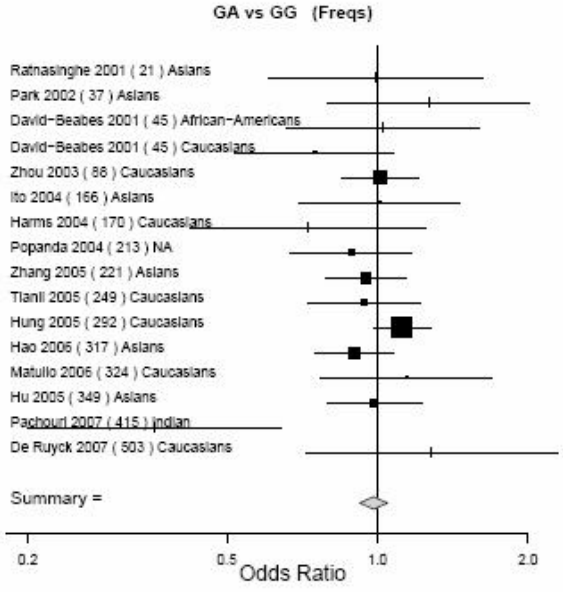
	Heterogenity test X	4.33				2.68				3.44			
	p-value	0.74				0.91				0.84			
XRCC1 cod 399 Lung NA NA NA	Forrest plot Funnel plot Radial plot												
		GA vs GG				AA vs GG				GA + AA vs GG			
		OR	95% CI		weights	OR	95% CI		weights	OR	95% CI		weights
Adjusted OR	David-Beabes 2001 (45) African-Americans	1.1	0.67	1.81	12.02	0.6	0.18	2.03	2.58	NA	NA	NA	NA
Random effects meta-analysis	David-Beabes 2001 (45) Caucasians	0.8	0.52	1.24	14.67	0.6	0.29	1.25	7.15	NA	NA	NA	NA
	De Ruyck 2007 (503) Caucasians	1.44	0.77	2.71	8.18	1.62	0.66	3.98	4.76	NA	NA	NA	NA
	Hao 2006 (317) Asians	0.89	0.74	1.07	36.83	0.86	0.62	1.19	37.1	0.88	0.74	1.05	28.92
	Harms 2004 (170) Caucasians	1.69	0.95	3	9.6	1.1	0.38	3.21	3.36	0.74	0.43	1.28	9.51
	Hu 2005 (349) Asians	0.98	0.78	1.23	31.44	0.83	0.55	1.26	21.81	0.95	0.77	1.18	26
	Hung 2005 (292) Caucasians	1.06	0.92	1.22	42.57	0.95	0.77	1.18	84.33	1.03	0.9	1.18	31.86
	Ito 2004 (166) Asians	1.02	0.69	1.5	17.38	1.35	0.65	2.8	7.24	NA	NA	NA	NA
	Matullo 2006 (324) Caucasians	1.37	0.83	2.26	12.04	0.5	0.2	1.23	4.74	NA	NA	NA	NA
	Pachouri 2007 (415) Indian	0.3	0.16	0.56	8.22	0.4	0.16	1.02	4.35	0.6	0.45	0.79	21.49
	Park 2002 (37) Asians	1.28	0.8	2.05	13.17	2.02	0.75	5.44	3.92	1.36	0.87	2.14	12.52
	Popanda 2004 (213) NA	0.88	0.64	1.2	22.72	0.97	0.62	1.52	19.11	NA	NA	NA	NA
	Ratnasinghe 2001 (21) Asians	1	0.61	1.63	12.36	1.3	0.49	3.44	4.06	1	0.61	1.63	11.21
	Zhang 2005 (221) Asians	0.93	0.77	1.13	35.87	1.17	0.85	1.61	37.66	NA	NA	NA	NA
	Zhou 2003 (88) Caucasians	1	0.78	1.27	29.72	1.3	0.97	1.74	44.48	NA	NA	NA	NA
	Ryk 2006 (303) Mixed	NA	NA	NA	NA	NA	NA	NA	NA	1.71	0.72	4.05	4.54
	Ryk 2006 (303) Mixed	NA	NA	NA	NA	NA	NA	NA	NA	0.61	0.36	1.03	10.15
	Ryk 2006 (303) Mixed	NA	NA	NA	NA	NA	NA	NA	NA	0.81	0.52	1.26	13.04
	Summary	0.99	0.88	1.1		1	0.89	1.12		0.89	0.77	1.04	
	Heterogenity test X	25.48				18.3				20.41			
	p-value	0.03				0.19				0.02			
Crude OR	David-Beabes 2001 (45) African-Americans	1.03	0.66	1.6	19.09	0.52	0.14	1.97	3.36	0.97	0.63	1.49	16.56
Fixed effects meta-analysis	David-Beabes 2001 (45) Caucasians	0.75	0.52	1.08	33.36	0.63	0.34	1.14	14.5	0.72	0.51	1.02	23.27
	De Ruyck 2007 (503) Caucasians	1.28	0.72	2.29	10.16	1.77	0.77	4.04	4.26	1.38	0.8	2.39	11.17
	Hao 2006 (317) Asians	0.9	0.75	1.08	124.81	0.83	0.61	1.14	43.24	0.89	0.75	1.05	51.98
	Harms 2004 (170) Caucasians	0.72	0.42	1.25	15.31	1.07	0.38	2.96	3.58	0.77	0.46	1.29	12.2
	Hu 2005 (349) Asians	0.99	0.79	1.23	81.12	0.81	0.54	1.22	25.67	0.96	0.78	1.18	43.46
	Hung 2005 (292) Caucasians	1.12	0.98	1.28	209.4	1.01	0.83	1.23	98.27	1.09	0.97	1.24	63.53
	Ito 2004 (166) Asians	1.01	0.7	1.46	28.26	1.39	0.7	2.77	6.52	1.06	0.75	1.5	22.99
	Matullo 2006 (324) Caucasians	1.14	0.77	1.7	22.87	0.52	0.23	1.17	9.74	1.01	0.69	1.49	19.84
	Pachouri 2007 (415) Indian	0.36	0.2	0.64	18.93	0.47	0.2	1.09	7.7	0.38	0.22	0.66	11.03
	Park 2002 (37) Asians	1.27	0.79	2.02	15.79	2.29	0.87	6.09	2.94	1.38	0.88	2.16	15.76
	Popanda 2004 (213) NA	0.89	0.67	1.17	51.82	0.87	0.58	1.29	26.1	0.88	0.68	1.15	33.29
	Ratnasinghe 2001 (21) Asians	1	0.61	1.63	15.85	1.38	0.5	3.8	3.07	1.04	0.65	1.67	14.33
	Tianli 2005 (249) Caucasians	0.94	0.72	1.22	59.21	0.83	0.56	1.25	26.31	0.92	0.72	1.17	36.55
	Zhang 2005 (221) Asians	0.95	0.79	1.14	112.38	1.14	0.84	1.55	37.88	0.98	0.83	1.17	50.66
	Zhou 2003 (88) Caucasians	1.01	0.85	1.2	125.48	1.29	0.99	1.67	50.71	1.07	0.91	1.26	53.43
	Summary	0.98	0.92	1.05		0.99	0.89	1.1		0.96	0.88	1.05	

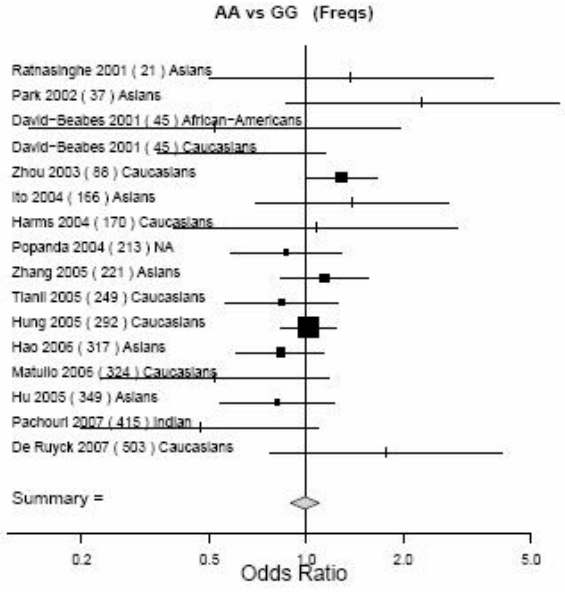
Internet

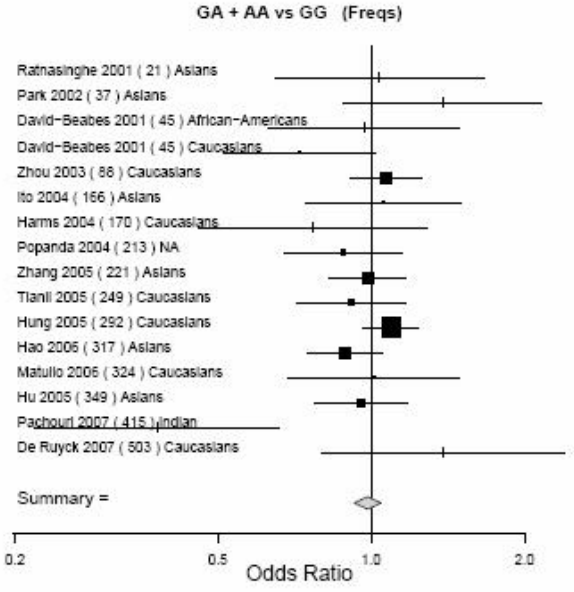

GA vs GG (ORs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
David-Beabes 2001 (45) African-Americans
David-Beabes 2001 (45) Caucasians
Zhou 2003 (86) Caucasians
Ito 2004 (166) Asians
Harms 2004 (170) Caucasians
Popanda 2004 (213) NA
Zhang 2005 (221) Asians
Hung 2005 (292) Caucasians
Hao 2006 (317) Asians
Matullo 2006 (324) Caucasians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
De Ruyck 2007 (503) Caucasians
Summary =
0.2
0.5
1.0
2.0
Odds Ratio

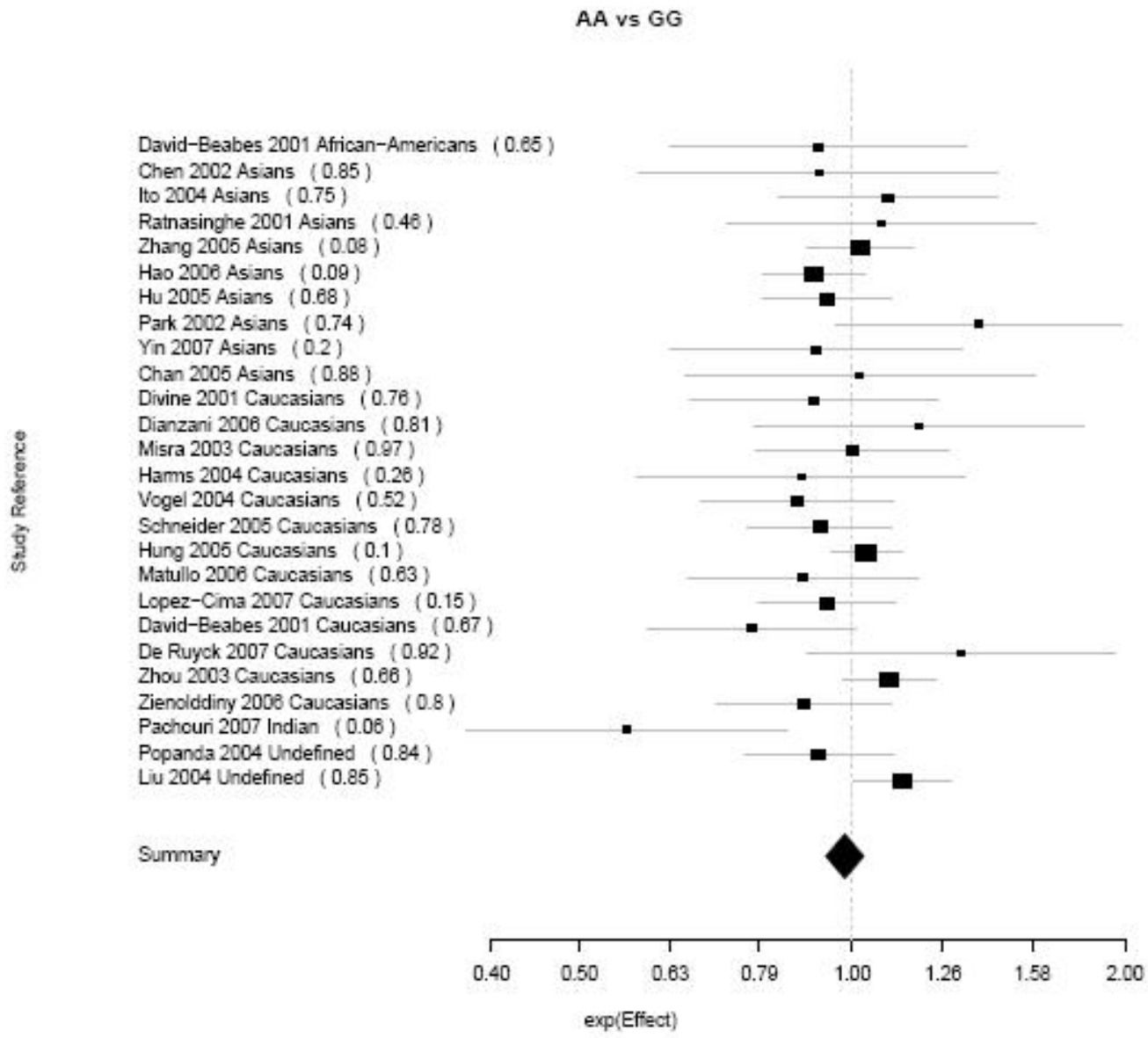

AA vs GG (ORs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
David-Beabes 2001 (45) African-Americans
David-Beabes 2001 (45) Caucasians
Zhou 2003 (86) Caucasians
Ito 2004 (166) Asians
Harms 2004 (170) Caucasians
Popanda 2004 (213) NA
Zhang 2005 (221) Asians
Hung 2005 (292) Caucasians
Hao 2006 (317) Asians
Matullo 2006 (324) Caucasians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
De Ruyck 2007 (503) Caucasians
Summary =
0.2
0.5
1.0
2.0
5.0
Odds Ratio


GA + AA vs GG (ORs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
Harms 2004 (170) Caucasians
Hung 2005 (292) Caucasians
Ryk 2006 (303) Mixed
Ryk 2006 (303) Mixed
Ryk 2006 (303) Mixed
Hao 2006 (317) Asians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
Summary =
0.5
1.0
2.0
Odds Ratio


GA vs GG (Freqs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
David-Beabes 2001 (45) African-Americans
David-Beabes 2001 (45) Caucasians
Zhou 2003 (86) Caucasians
Ito 2004 (166) Asians
Harms 2004 (170) Caucasians
Popanda 2004 (213) NA
Zhang 2005 (221) Asians
Tianli 2005 (249) Caucasians
Hung 2005 (292) Caucasians
Hao 2006 (317) Asians
Matullo 2006 (324) Caucasians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
De Ruyck 2007 (503) Caucasians
Summary =
0.2
0.5
1.0
2.0
Odds Ratio


AA vs GG (Freqs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
David-Beabes 2001 (45) African-Americans
David-Beabes 2001 (45) Caucasians
Zhou 2003 (86) Caucasians
Ito 2004 (166) Asians
Harms 2004 (170) Caucasians
Popanda 2004 (213) NA
Zhang 2005 (221) Asians
Tianli 2005 (249) Caucasians
Hung 2005 (292) Caucasians
Hao 2006 (317) Asians
Matullo 2006 (324) Caucasians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
De Ruyck 2007 (503) Caucasians
Summary =
0.2
0.5
1.0
2.0
5.0
Odds Ratio


GA + AA vs GG (Freqs)
Ratnasinghe 2001 (21) Asians
Park 2002 (37) Asians
David-Beabes 2001 (45) African-Americans
David-Beabes 2001 (45) Caucasians
Zhou 2003 (86) Caucasians
Ito 2004 (166) Asians
Harms 2004 (170) Caucasians
Popanda 2004 (213) NA
Zhang 2005 (221) Asians
Tianli 2005 (249) Caucasians
Hung 2005 (292) Caucasians
Hao 2006 (317) Asians
Matullo 2006 (324) Caucasians
Hu 2005 (349) Asians
Pachouri 2007 (415) Indian
De Ruyck 2007 (503) Caucasians
Summary =
0.2
0.5
1.0
2.0
Odds Ratio


AA vs GG
Study Reference
David−Beabes 2001 African−Americans (0.65)
Chen 2002 Asians (0.85)
Ito 2004 Asians (0.75)
Ratnasinghe 2001 Asians (0.46)
Zhang 2005 Asians (0.08)
Hao 2006 Asians (0.09)
Hu 2005 Asians (0.68)
Park 2002 Asians (0.74)
Yin 2007 Asians (0.2)
Chan 2005 Asians (0.88)
Divine 2001 Caucasians (0.76)
Dianzani 2006 Caucasians (0.81)
Misra 2003 Caucasians (0.97)
Harms 2004 Caucasians (0.26)
Vogel 2004 Caucasians (0.52)
Schneider 2005 Caucasians (0.78)
Hung 2005 Caucasians (0.1)
Matullo 2006 Caucasians (0.63)
Lopez−Cima 2007 Caucasians (0.15)
David−Beabes 2001 Caucasians (0.67)
De Ruyck 2007 Caucasians (0.92)
Zhou 2003 Caucasians (0.66)
Zienolddiny 2006 Caucasians (0.8)
Pachouri 2007 Indian (0.06)
Popanda 2004 Undefined (0.84)
Liu 2004 Undefined (0.85)
Summary
0.40
0.50
0.63
0.79
1.00
1.26
1.58
2.00
exp(Effect)

Relevance

DNA repair is a field in which genotyping data in relation to cancer have increased exponentially in recent years.

Environmental carcinogens, such as polycyclic aromatic hydrocarbons, aromatic amines or N-nitroso compounds, predominantly form DNA adducts but also generate interstrand cross-links and reactive oxygen species [which induce base damage, abasic sites, and single and double strand breaks (SSBs and DSBs)].

Unrepaired damage can result in apoptosis or may lead to unregulated cell growth and cancer.

Meta-analyses were performed on 213 associations (1083 studies) between specific variants and specific cancer types that had been tested by at least 2 independent studies.

A total of 14 nominally significant associations were recorded for 8 genes in the main (allele-based) analyses (*TP53, BRCA2, ERCC2, ERCC5, XPA, XRCC1, XRCC2, XRCC3*).

We graded the epidemiological strength of each of these nominal associations with criteria that assess amount of evidence, replication, and protection from bias.

Overview:

* main analyses (OR per allele)

* secondary analyses:
dominant vs recessive models
by smoking, histological type, gender, ethnicity, study design

* heterogeneity (I^2) (values of 50% or higher are considered to express large between-study heterogeneity and values of 25-50% express moderate between-study heterogeneity)

* bias: exclusion of first study, exclusion of HWE-violating studies, small-study effect bias (modified Egger test), observed vs expected statistical significant results (Ioannidis and Trikalinos test)

* overall grading

We applied the Venice criteria:

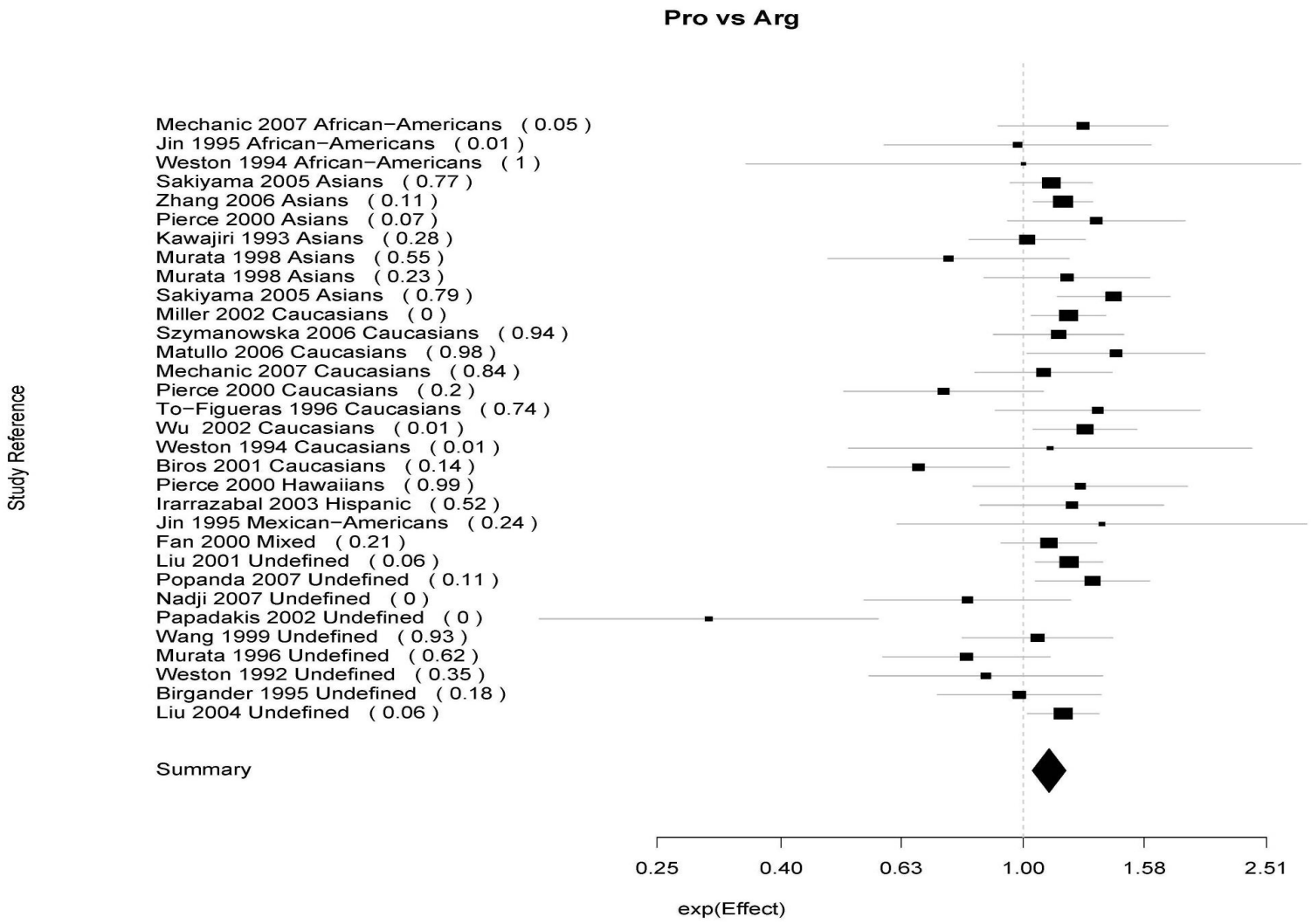
Grades A, B and C for amount of evidence, replication and protection from bias

Associations that get three A grades are considered to have "strong" epidemiological credibility, associations that get any B but not any C grade are assigned "moderate" credibility, and associations that get any C grade as considered to have "weak" credibility.

Provisional results

Nominally significant associations

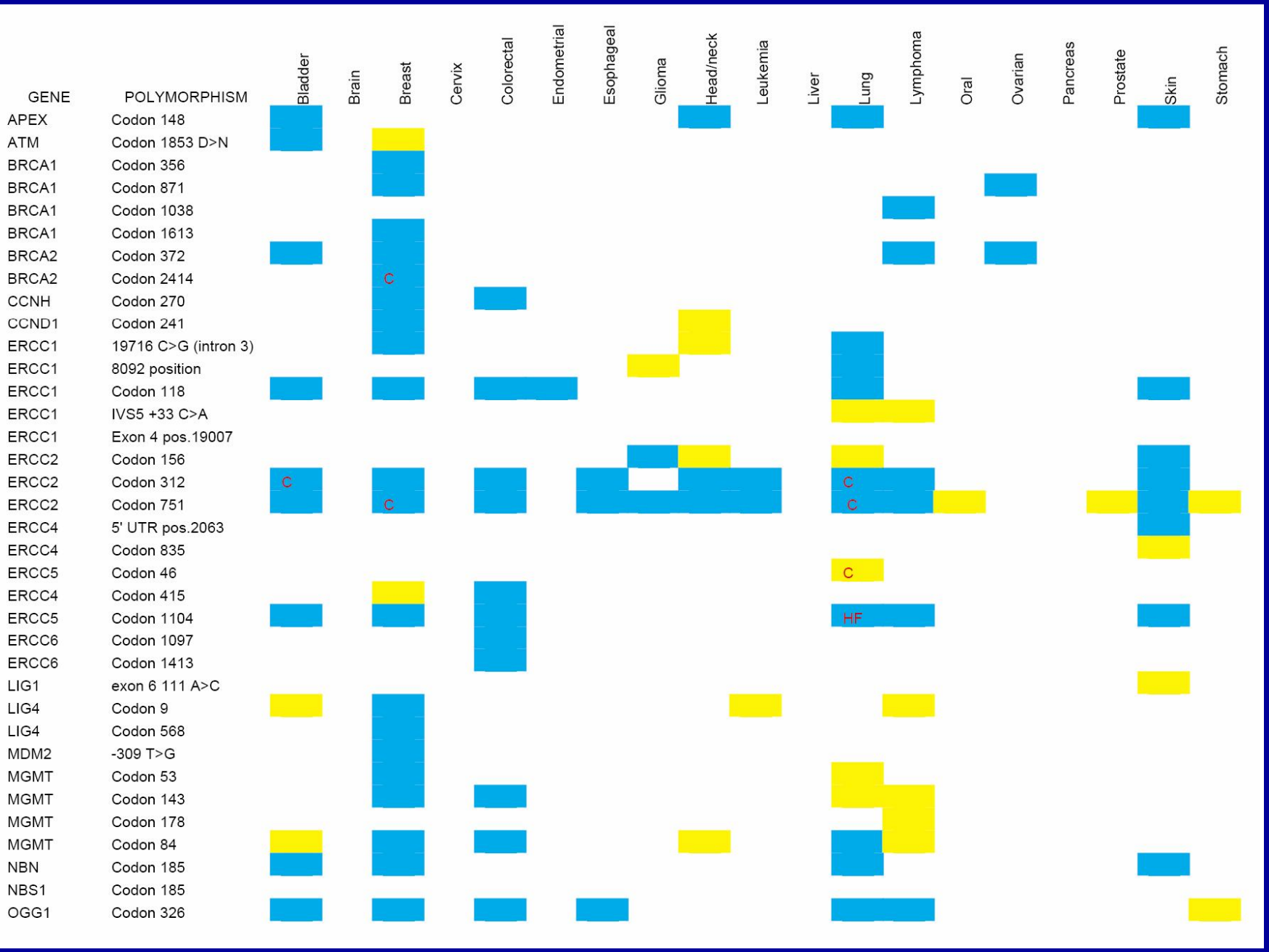
Gene		Cancer	Studies	N*	OR (95% CI)	P-value
BRCA2	cod. 2414	Breast	3	16,344	1.06 (1.01-1.12	0.03
ERCC2	Cod.312	Bladder	4	4,006	1.12 (1.01-1.24)	0.03
ERCC2	Cod.751	Breast	15	11,660	1.06 (1.01-1.12)	0.02
ERCC2	cod 312	Lung	14	920	1.07 (1.01-1.15)	0.02
ERCC2	Cod.751	Lung	19	14,989	1.13 (1.04-1.22)	0.003
ERCC5	cod 46	Lung	2	21,477	0.73 (0.61-0.89)	0.001
TP53	Cod 72	Lung	32	4,843	1.11 (1.04-1.18)	0.001
TP53	intron 6	Breast	6	3,779	1.08 (1.01-1.16)	0.03
XPA	23 G>A	Headneck	3	5,918	0.91 (0.84-0.99)	0.03
XRCC1	-77T>C	Lung	3	1,555	1.39 (1.20-1.61)	10^{-5}
XRCC1	cod 194	Skin	3	16,344	0.75 (0.57-0.99)	0.04
XRCC1	cod 194	Stomach	5	4,006	0.82 (0.70-0.97)	0.02
XRCC2	cod 188	Colorectal	2	11,660	1.14 (1.00-1.30)	0.05
XRCC3	Cod.241	Colorectal	5	12,525	0.89 (0.81-0.97)	0.01

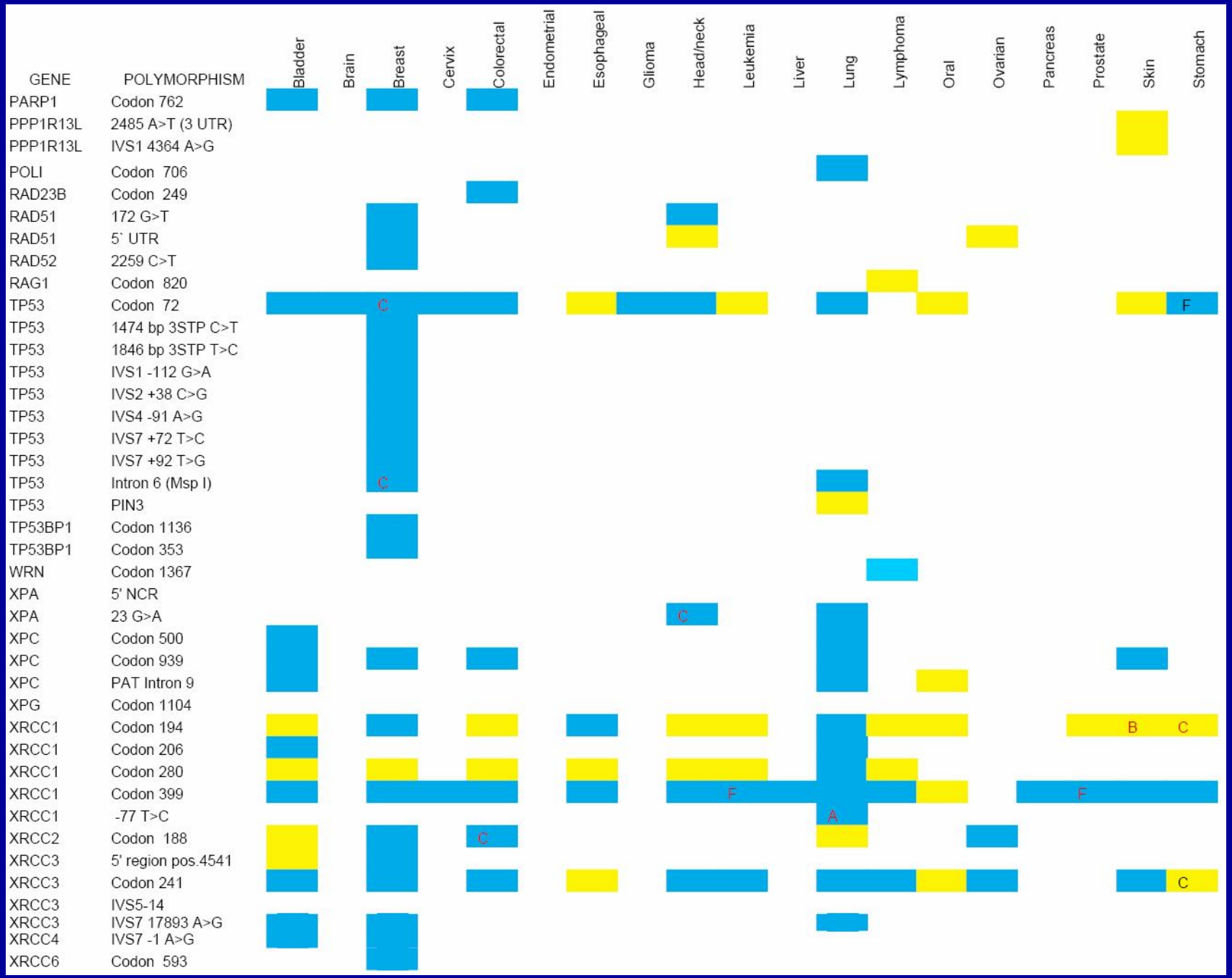

Pro vs Arg
Study Reference
Mechanic 2007 African−Americans (0.05)
Jin 1995 African−Americans (0.01)
Weston 1994 African−Americans (1)
Sakiyama 2005 Asians (0.77)
Zhang 2006 Asians (0.11)
Pierce 2000 Asians (0.07)
Kawajiri 1993 Asians (0.28)
Murata 1998 Asians (0.55)
Murata 1998 Asians (0.23)
Sakiyama 2005 Asians (0.79)
Miller 2002 Caucasians (0)
Szymanowska 2006 Caucasians (0.94)
Matullo 2006 Caucasians (0.98)
Mechanic 2007 Caucasians (0.84)
Pierce 2000 Caucasians (0.2)
To−Figueras 1996 Caucasians (0.74)
Wu 2002 Caucasians (0.01)
Weston 1994 Caucasians (0.01)
Biros 2001 Caucasians (0.14)
Pierce 2000 Hawaiians (0.99)
Irarrazabal 2003 Hispanic (0.52)
Jin 1995 Mexican−Americans (0.24)
Fan 2000 Mixed (0.21)
Liu 2001 Undefined (0.06)
Popanda 2007 Undefined (0.11)
Nadji 2007 Undefined (0)
Papadakis 2002 Undefined (0)
Wang 1999 Undefined (0.93)
Murata 1996 Undefined (0.62)
Weston 1992 Undefined (0.35)
Birgander 1995 Undefined (0.18)
Liu 2004 Undefined (0.06)
Summary
0.25
0.40
0.63
1.00
1.58
2.51
exp(Effect)

Gene		Cancer	Minor N	I^2	First-out
BRCA2	cod. 2414	Breast	7,153	0	na
ERCC2	Cod.312	Bladder	2,855	0	1.08 (0.87-1.34)
ERCC2	Cod.751	Breast	8,964	0	1.07 (1.01-1.13)
ERCC2	cod 312	Lung	6,365	0	1.08 (1.01-1.16)
ERCC2	Cod.751	Lung	8,303	41	1.11 (1.04-1.18)
ERCC5	cod 46	Lung	905	0	na
TP53	Cod 72	Lung	13,923	44	1.11 (1.04-1.18)
TP53	intron 6	Breast	24,739	3	na
XPA	23 G>A	Head/neck	3,237	3	na
XRCC1	-77T>C	Lung	1,020	11	1.30 (1.01-1.66)
XRCC1	cod 194	Skin	389	6	na
XRCC1	cod 194	Stomach	913	0	0.85 (0.71-1.02)
XRCC2	cod 188	Colorectal	862	0	na
XRCC3	Cod.241	Colorectal	3,106	0	0.90 (0.80-1.01)

Venice grading

Gene		Cancer	Amount	Replication	Bias	Reason	ALL
BRCA2	Cod2414	Breast	A	A	C	Low OR, F	C
ERCC2	Codon312	Bladder	A	A	C	Low OR, F, R, HWE	C
ERCC2	Codon751	Breast	A	A	C	Low OR	C
ERCC2	Codon 312	Lung	A	A	C	Low OR	C
ERCC2	Codon751	Lung	A	B	C	Low OR	C
ERCC5	Codon 46	Lung	B	A	C	F, HWE	C
TP53	Codon 72	Lung	A	B	C	Low OR, R	C
TP53	intron 6	Breast	A	A	C	Low OR, F, HWE	C
XPA	23 G>A	Head/neck	A	A	C	Low OR, F	C
XRCC1	-77T>C	Lung	A	A	A	-	A
XRCC1	Codon 194	Skin	B	A	A	-	B
XRCC1	Codon 194	Stomach	B	A	C	F	C
XRCC2	Codon 188	Colorectal	B	A	C	Low OR, F, HWE	C
XRCC3	Codon 241	Colorectal	A	A	C	Low OR	C


GENE
POLYMORPHISM
Bladder
Brain
Breast
Cervix
Colorectal
Endometrial
Esophageal
Glioma
Head/neck
Leukemia
Liver
Lung
Lymphoma
Oral
Ovarian
Pancreas
Prostate
Skin
Stomach
APEX Codon 148
ATM Codon 1853 D>N
BRCA1 Codon 356
BRCA1 Codon 871
BRCA1 Codon 1038
BRCA1 Codon 1613
BRCA2 Codon 372
BRCA2 Codon 2414
CCNH Codon 270
CCND1 Codon 241
ERCC1 19716 C>G (intron 3)
ERCC1 8092 position
ERCC1 Codon 118
ERCC1 IVS5 +33 C>A
ERCC1 Exon 4 pos.19007
ERCC2 Codon 156
ERCC2 Codon 312
ERCC2 Codon 751
ERCC4 5' UTR pos.2063
ERCC4 Codon 835
ERCC5 Codon 46
ERCC4 Codon 415
ERCC5 Codon 1104
ERCC6 Codon 1097
ERCC6 Codon 1413
LIG1 exon 6 111 A>C
LIG4 Codon 9
LIG4 Codon 568
MDM2 -309 T>G
MGMT Codon 53
MGMT Codon 143
MGMT Codon 178
MGMT Codon 84
NBN Codon 185
NBS1 Codon 185
OGG1 Codon 326
C
C
C
C
C
C
C
HF


GENE
POLYMORPHISM
Bladder
Brain
Breast
Cervix
Colorectal
Endometrial
Esophageal
Glioma
Head/neck
Leukemia
Liver
Lung
Lymphoma
Oral
Ovarian
Pancreas
Prostate
Skin
Stomach
PARP1 Codon 762
PPP1R13L 2485 A>T (3 UTR)
PPP1R13L IVS1 4364 A>G
POLI Codon 706
RAD23B Codon 249
RAD51 172 G>T
RAD51 5` UTR
RAD52 2259 C>T
RAG1 Codon 820
TP53 Codon 72
TP53 1474 bp 3STP C>T
TP53 1846 bp 3STP T>C
TP53 IVS1 -112 G>A
TP53 IVS2 +38 C>G
TP53 IVS4 -91 A>G
TP53 IVS7 +72 T>C
TP53 IVS7 +92 T>G
TP53 Intron 6 (Msp I)
TP53 PIN3
TP53BP1 Codon 1136
TP53BP1 Codon 353
WRN Codon 1367
XPA 5' NCR
XPA 23 G>A
XPC Codon 500
XPC Codon 939
XPC PAT Intron 9
XPG Codon 1104
XRCC1 Codon 194
XRCC1 Codon 206
XRCC1 Codon 280
XRCC1 Codon 399
XRCC1 -77 T>C
XRCC2 Codon 188
XRCC3 5' region pos.4541
XRCC3 Codon 241
XRCC3 IVS5-14
XRCC3 IVS7 17893 A>G
XRCC4 IVS7 -1 A>G
XRCC6 Codon 593
C
C
C
F
C
F
A
F
B
C
C

Comments

Surprisingly large amount of evidence on selected associations

Racial descent
Most studies in Caucasians, difficult to assess ethnicity

TP53 lung:
Afro-Americans 1.15 (0.89-1.50)
Asians 1.15 (1.04-1.27)
Caucasians 1.10 (0.95-1.26)

Types of participants
When a comparison between hospital-based and population-based studies was feasible, results were consistent

ERCC2 312, lung HB 1.08, P-B 1.08
ERCC2 751, breast H-B 1.02 P-B 1.07
ERCC2 751, lung H-B 1.13 P-B 1.02

Difficult and open issues:

1. Biological plausibility (little is known on functional significance: genotype-phenotype correlation project led by Matullo within ECNIS)

2. Pleiotropism, i.e. one can expect associations with several types of cancer (lack of specificity). ERCC2 and XRCC1 involved with several cancer types, BER and NER more important than e.g. MGMT

3. Effect of gene/SNP combination and analytical strategies

DNA repair ECNIS projects

1) **Use of lymphoblastoid cell lines for the identification of at risk genotypes: validation of cell lines for the DNA repair response**
(P.I. Matullo ISI; UC)
   - DNA repair gene expression (ERCC1, OGG1)
   - RO19-8022 BER modified Comet for cell extracts

2) **Development and validation of phenotypical nucleotide excision DNA repair (NER) assays for use in molecular epidemiology**
(P.I. Kirsch-Volders VUB; UM; ISI, NHRF, UO)
   - BPDE NER modified Comet for cell extracts
   - Comparing methods for preserving cell extracts
   - Development of a MMR phenotypic assay

2) **Assessment and reduction of comet assay variation in relation to DNA damage and DNA repair phenotype**
(P.I. Moller UC; KI; UM, ULEIC)

EXTENSIVE DNA REPAIR GENOTYPE-PHENOTYPE CORRELATION: AN ITALIAN NETWORK

- P.I. Matullo (University of Turin, ISS Rome, CSPO Florence)
- Cryopreserved lymphocytes from 300 healthy subjects
- 1536 selected DNA repair SNPs (Illumina's chip)
- DNA repair phenotipic assays on lymphocytes after exposure to low dose of carcinogens:
  - BER modified COMET for cell extracts (RO19)
  - COMET incision assay (BPDE)
  - Micronuclei and CA (γ rays)
  - Phosphorylation of hystone H2AX (γ rays)
  - 32P-postlabelling DNA adducts (BPDE)
  - BER DNA glycosylases activities (MMR?)
  - DNA repair gene expression

Additive combinations of multiple weak risk factors (population attributable fraction)

A total of 100 true risk factors were simulated, each conferring a relative risk of 1.25, and with a MAF of 10%. In each line are counted the subjects carrying combinations of any k risk alleles.
OR = odds ratios, calculated with a simple additive model. For example, for subjects with 10 risk alleles the relative risk would be 3.5. These subjects would represent 13% of the population, and over 54% of the population would carry 10 risk alleles or more.

(from Vineis P et al, Expectations and challenges stemming from genome-wide association studies, submitted to CEBP)

Number of risk factors (k)	% of subjects who carry k risk alleles	% of subjects who carry at least k risk alleles	OR	PAF
0	0.003%	-	1.00	-
1	0.030%	99.99%	1.25	0.0001
2	0.162%	99.97%	1.50	0.0008
3	0.589%	99.81%	1.75	0.0044
4	1.587%	99.22%	2.00	0.0156
5	3.387%	97.63%	2.25	0.0406
6	5.958%	94.24%	2.50	0.0820
7	8.890%	88.28%	2.75	0.1346
8	11.482%	79.39%	3.00	0.1868
9	13.042%	67.91%	3.25	0.2269
10	13.187%	54.87%	3.50	0.2479
11	11.988%	41.68%	3.75	0.2479
12	9.879%	29.70%	4.00	0.2286
13	7.430%	19.82%	4.25	0.1945
14	5.130%	12.39%	4.50	0.1522
15	3.268%	7.26%	4.75	0.1092
16	1.929%	3.99%	5.00	0.0716
17	1.059%	2.06%	5.25	0.0431
18	0.543%	1.00%	5.50	0.0238
19	0.260%	0.46%	5.75	0.0122
20	0.117%	0.20%	6.00	0.0058
100	10^{-98}%	5.38×10^{-13}%	26.00	0.0000

Analytical strategies

Summary of some important differences between modelling for prediction and modelling for understanding.

	Prediction	Understanding
Question	Who is at high risk?	Which genes interact?
Consequences	Treatment of Subjects	Bioinformatics, replication studies
Model Choice	Less important	Vital
Key model terms	Joint effects	Interactions
Model complexity	May include high order terms	Low order terms

Methods to model gene-gene interactions

Extensions to regression analysis

Automated Detection of Informative Combined Effects (DICE)(Tahri-Daizadeh N et al. *Genome Research* 13: 1952-1960, 2003)

Classification and Regression Trees (CART)/ Patterning and Recursive Partitioning (PRP)(Bastone L et al. *Human Heredity* 58: 82-92, 2004)

Logic Regression (Kooperberg et al. *Genetic Epidemiology* (S1): 626-631, 2001)

Penalized Logistic Regression (Zhu et al. *Biostatistics* 5: 427 – 443, 2004)

Multivariate adaptive regression spline (Cook et al. Statistics in Medicine 23: 1439 – 1453, 2004)

Data reduction approaches

Combinatorial Partitioning Method (CPM)(Nelson et al., *Genome Research* 11:458-70, 2001)

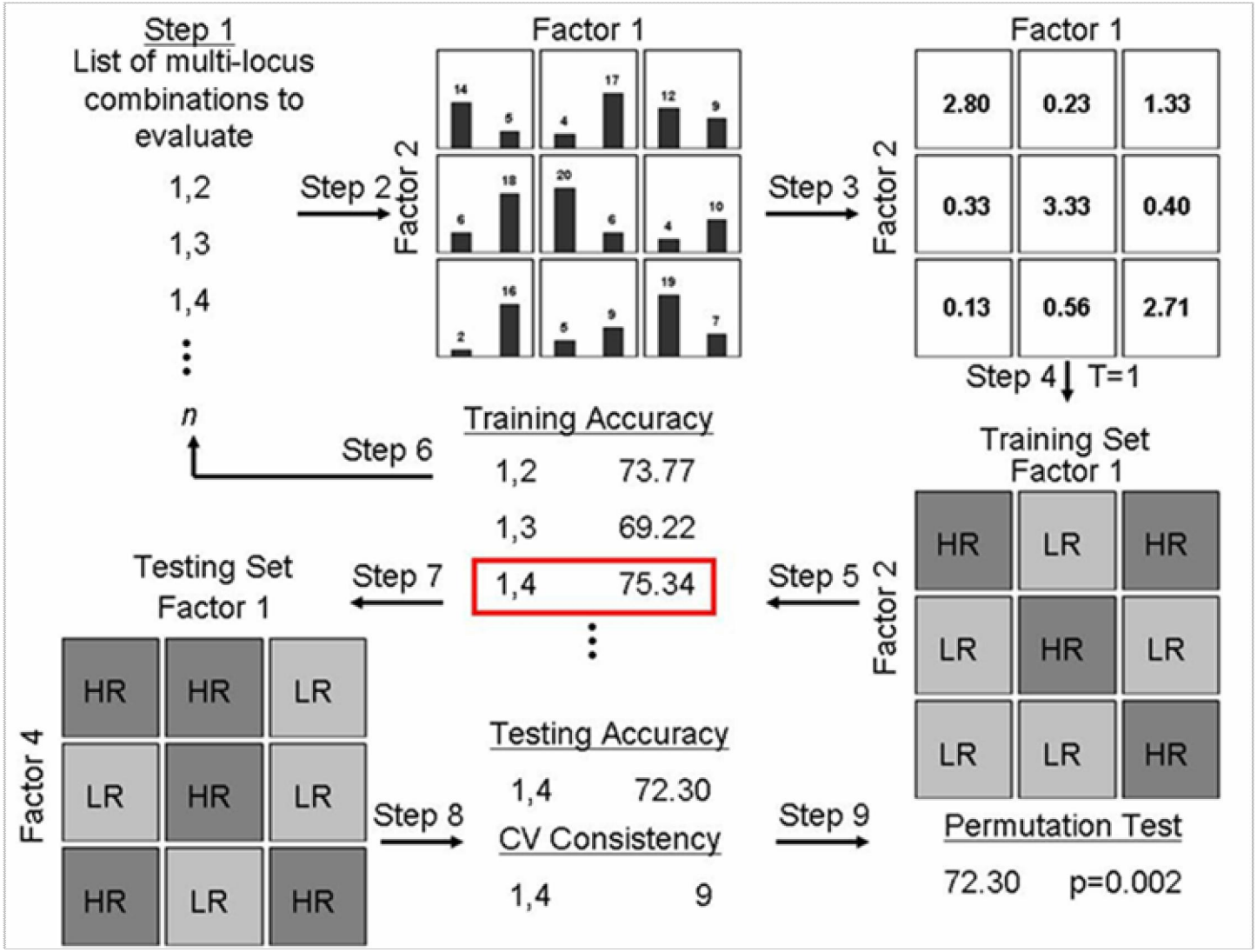
Restricted Partitioning Method (RPM)(Culverhouse R, Klein T, Shannon W. *Genet Epidemiol* 27:141-52, 2004)

Multifactor Dimensionality Reduction (MDR)(Ritchie et al., *American Journal of Human Genetics* 69:138-147, 2001)

Set Association (Hoh J, Wille A, and Ott J. *Genome Research* 11: 2115-2119, 2001).

Multifactor Dimensionality Reduction

Ritchie et al., *American Journal of Human Genetics* (2001)
Moore et al., *Journal of Theoretical Biology* (2006)



The End