Breast Cancer Susceptibility and HLA Complex: Source of Missing Heritability?

M. Tevfik DORAK

Environmental & Occupational Health
College of Public Health
Florida International University
Miami, USA
http://www.dorak.info

Cancer Institute of New Jersey
December 7, 2011

FIU FLORIDA INTERNATIONAL UNIVERSITY

Outline

Breast Cancer Susceptibility & GWAS

GWAS & Missing Heritability

GWAS Myths

HLA & Breast Cancer Susceptibility

Breast Cancer Susceptibility

Established and probable risk factors for breast cancer

Factor	Relative risk	High risk group
Age	>10	Elderly
Geographical location	5	Developed country
Age at menarche	3	Menarche before age 11
Age at menopause	2	Menopause after age 54
Age at first full pregnancy	3	First child in early 40s
Family history	≥2	Breast cancer in first degree relative when young
Previous benign disease	4-5	Atypical hyperplasia
Cancer in other breast	>4	
Socioeconomic group	2	Groups I and II
Diet	1.5	High intake of saturated fat
Body weight:		
Premenopausal	0.7	Body mass index >35
Postmenopausal	2	Body mass index >35
Alcohol consumption	1.3	Excessive intake
Exposure to ionising radiation	3	Abnormal exposure in young females after age 10
Taking exogenous hormones:		
Oral contraceptives	1.24	Current use
Hormone replacement therapy	1.35	Use for ≥10 years
Diethylstilbestrol	2	Use during pregnancy

FIU | FLORIDA INTERNATIONAL UNIVERSITY

K McPherson, CM Steel, JM Dixon. *BMJ* 2000;321(9):624-8

Breast Cancer Susceptibility

Risk Calculator

(Click a question number for a brief explanation, or read all explanations.)

1. Does the woman have a medical history of any breast cancer or of ductal carcinoma in situ (DCIS) or lobular carcinoma in situ (LCIS)? — Select
2. What is the woman's age? *This tool only calculates risk for women 35 years of age or older.* — Select
3. What was the woman's age at the time of her first menstrual period? — Select
4. What was the woman's age at the time of her first live birth of a child? — Select
5. How many of the woman's first-degree relatives - mother, sisters, daughters - have had breast cancer? — Select
6. Has the woman ever had a breast biopsy? — Select
   - 6a. How many breast biopsies (positive or negative) has the woman had? — Select
   - 6b. Has the woman had at least one breast biopsy with atypical hyperplasia? — Select
7. What is the woman's race/ethnicity? — Select

Calculate Risk >

Breast Cancer Susceptibility & Genetics

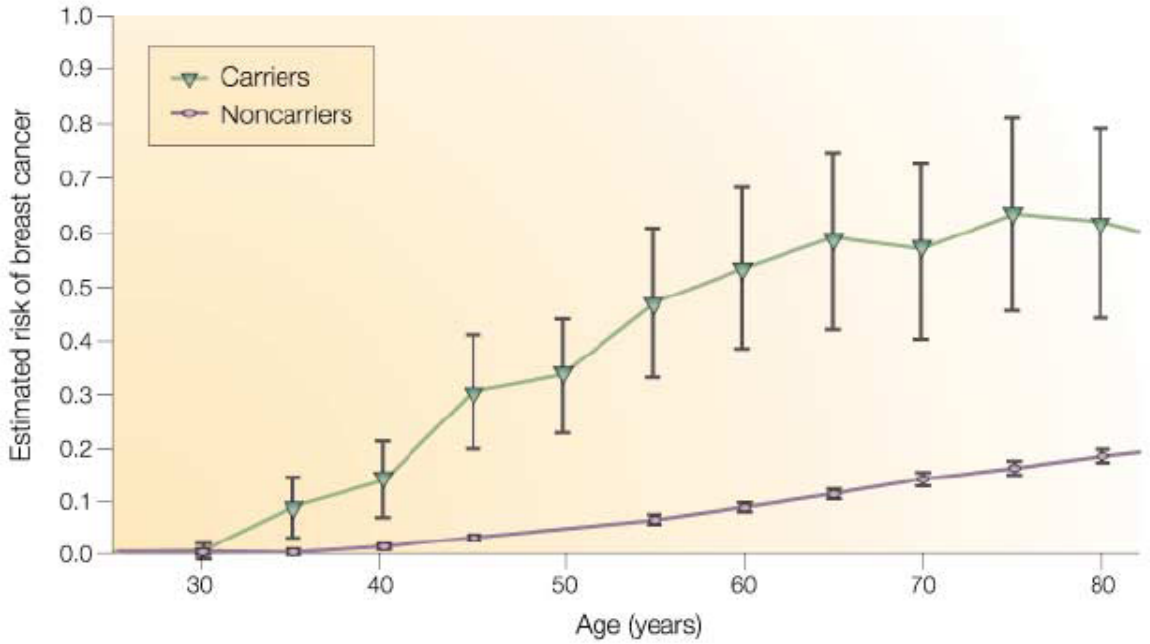


Figure 2 | **Risk of breast cancer caused by *BRCA1* and *BRCA2* mutations.** Risks for carriers versus non-carriers are shown. The curves show the estimated cumulative lifetime risk by a particular age, with 95% confidence intervals. Although cumulative risk does not decrease with age, the estimation method allows declines, which explains the decrease in cumulative risk for carriers in the 65–70 and 75–80 age intervals[15]. The data represented are from 5,318 Ashkenazi Jews. All individuals were genotyped for two specific mutations in each of *BRCA1* and *BRCA2*, yielding 120 mutants, which is 2.3% of the sample. There was no significant difference in risk between carriers of *BRCA1* and *BRCA2* mutations. Modified with permission from REF. 15 © (1997) Massachussets Medical Society.

Cumulatively, no more than 60% of women with a BRCA1/2 mutation will develop breast cancer by age 80.

Frank SA. *Nat Rev Genetics* 2004;5:764-72

Breast Cancer Susceptibility & Genetics

Box 1 Classes and key features of known breast cancer susceptibility alleles

High-penetrance breast cancer susceptibility genes

Examples: *BRCA1*, *BRCA2*, *TP53*

- **Risk variants:** Multiple, different mutations that predominantly cause protein truncation
- **Frequency:** Rare (population carrier frequency ≤0.1%)
- **Risk of breast cancer:** 10- to 20-fold relative risk
- **Primary strategy for identification:** Genome-wide linkage and positional cloning

Moderate-penetrance breast cancer susceptibility genes

Examples: *ATM*, *BRIP1*, *CHEK2*, *PALB2*

- **Risk variants:** Multiple, different mutations that predominantly cause protein truncation
- **Frequency:** Rare (population carrier frequency ≤0.6%)
- **Risk of breast cancer:** two- to fourfold relative risk
- **Primary strategy for identification:** Direct interrogation of candidate genes for coding variants in large, genetically enriched breast cancer case series and controls

Low-penetrance breast cancer susceptibility alleles

Examples: rs2981582 (*FGFR2*, 10q), rs3803662 (*TNRC9* (recently renamed *TOX3*), 16q), rs889312 (*MAP3K1*, 5q), rs3817198 (*LSP1*, 11p), rs13281615 (8q), rs13387042 (2q), rs1045485 (*CASP8*_D302H)

- **Risk variants:** Single-nucleotide polymorphisms that are causal or in linkage disequilibrium with the causal variant(s). May occur in noncoding, nongenic regions.
- **Frequency:** Common (population frequency 5–50%)
- **Risk of breast cancer:** up to ~1.25-fold (heterozygous) or 1.65-fold (homozygous) relative risk
- **Primary strategy for identification:** Genome-wide association studies of hundreds of thousands of SNPs in large breast cancer case-control series

Stratton & Rahman. *Nat Genet* 2008;40(1):17-21

Breast Cancer Susceptibility & GWAS

Table 1. Single-nucleotide polymorphisms associated with risk for breast cancer*

Chromosome locus	Candidate genes	SNPs	Per-allele ORs (95% CIs)	MAF	Reference
1p11.2	Pericentric *NOTCH2, FCGR1B*	rs11249433†	1.16 (1.09 to 1.24)	0.39	1
2q35	Intergenic	rs13387042†	1.20 (1.14 to 1.26)	0.50	2
		rs13387042	1.12 (1.09 to 1.15)	0.51	14
3p24.1	*SLC4A7, NEK10*	rs4973768	1.11 (1.08 to 1.13)	0.46	3
5q11.2	*MAP3K1, MIER3, C5orf35*	rs889312	1.13 (1.10 to 1.16)	0.28	4,6
		rs4415084†	1.23 (1.16 to 1.30)	0.44	5
		rs10941679†	1.27 (1.19 to 1.35)	0.29	5
5p12	Intergenic *MRPS30*	rs7703618	1.13 (1.08 to 1.18)	0.37	5
6q22.33	*RNF146, ECHDC1*	rs2180341	1.41 (1.25 to 1.59)	0.21	7‡
6q25.1	*ESR1*	rs2046210	1.29 (1.21 to 1.37)	0.35	8§
8q24	Intergenic	rs13281615†	1.08 (1.05 to 1.11)	0.40	4
10q26	*FGFR2*	rs2981582	1.26 (1.23 to 1.30)	0.38	4
		rs1219648	1.20 (1.07 to 1.42)	0.39	9
		rs1078806	1.26 (1.13 to 1.40)	0.40	7‡
11p15.5	*LSP1*	rs3817198	1.07 (1.04 to 1.11)	0.30	4
14q24.1	*RAD51L1*	rs999737	1.06 (1.01 to 1.14)	0.76	1
16p13.3	*A2BP1(FOX1)*	rs7203563	1.32 (1.11 to 1.57)	0.11	7‡
16q12	*TNRC9 (TOX3) LOC643714*	rs3803662	1.20 (1.16 to 1.24)	0.25	4
			1.28 (1.21 to 1.35)	0.27	2
17q23	*STXBP4*	rs6504950†	1.05 (1.03 to 1.09)	0.27	3

Breast Cancer Single-Nucleotide Polymorphisms: Statistical Significance and Clinical Utility

Kenneth Offit

Breast Cancer Susceptibility & GWAS

Table 3. Interactions between SNPs and established risk factors*

SNP	Gene	Chr	Location, bp‡
rs11249433	*NOTCH2*	1	120,982,136
rs1045485	*CASP8*	2	201,857,834
rs13387042	Intergenic	2	217,614,077
rs4973768	*SLC4A7*	3	27,391,017
rs4415084***	Intergenic	5	44,698,272
rs10941679	Intergenic	5	44,742,255
rs889312	*MAP3K1*	5	56,067,641
rs2180341	*RNF146*	6	127,642,323
rs2046210	Intergenic	6	151,990,059
rs13281615	Intergenic	8	128,424,801
rs2981582	*FGFR2*	10	123,342,308
rs3750817	*FGFR2*	10	123,322,567
rs3817198	*LSP1*	11	1,865,583
rs999737†††	*RAD51L1*	14	68,104,435
rs3803662	*TNRC9*	16	51,143,843
rs2075555	*COL1A1*	17	45,629,290
rs6504950	*COX11*	17	50,411,470

Interactions Between Genetic Variants and Breast Cancer Risk Factors in the Breast and Prostate Cancer Cohort Consortium

Daniele Campa, Rudolf Kaaks, Loïc Le Marchand, Christopher A. Haiman, Ruth C. Travis, Christine D. Berg, Julie E. Buring, Stephen J. Chanock, W. Ryan Diver, Lucie Dostal, Agnes Fournier, Susan E. Hankinson, Brian E. Henderson, Robert N. Hoover, Claudine Isaacs, Mattias Johansson, Laurence N. Kolonel, Peter Kraft, I-Min Lee, Catherine A. McCarty, Kim Overvad, Salvatore Panico, Petra H.M. Peeters, Elio Riboli, María José Sánchez, Fredrick R. Schumacher, Guri Skeie, Daniel O. Stram, Michael J. Thun, Dimitrios Trichopoulos, Shumin Zhang, Regina G. Ziegler, David J. Hunter, Sara Lindström, Federico Canzian

Manuscript received August 23, 2010; revised June 14, 2011; accepted June 21, 2011.

J Natl Cancer Inst 2011;103:1252–1263

Breast Cancer Susceptibility & GWAS

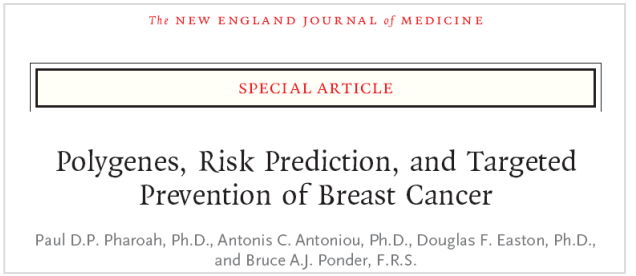
Table 1. Established Common Breast-Cancer Susceptibility Alleles.*

dbSNP No.	Gene†	Chromosome	Risk-Allele Frequency‡	Relative Risk per Allele‡	Fraction of Total Variance in Risk Explained§	Population Attributable Risk§	Study
					%		
rs2981582	*FGFR2*	10q	0.38	1.26	1.7	19	Easton et al.,[26] Hunter et al.[27]
rs3803662	*TNRC9, LOC643714*	16q	0.25	1.20	0.9	10	Easton et al.[26]
rs889312	*MAP3K1*	5q	0.28	1.13	0.4	7	Easton et al.[26]
rs3817198	*LSP1*	11p	0.30	1.07	0.1	4	Easton et al.[26]
rs13281615	None known	8q	0.40	1.08	0.2	6	Easton et al.[26]
rs13387042	None known	2q	0.50	1.20	1.2	19	Stacey et al.[28]
rs1053485	*CASP8*	2q	0.86	1.13	0.3	20	Cox et al.[25]

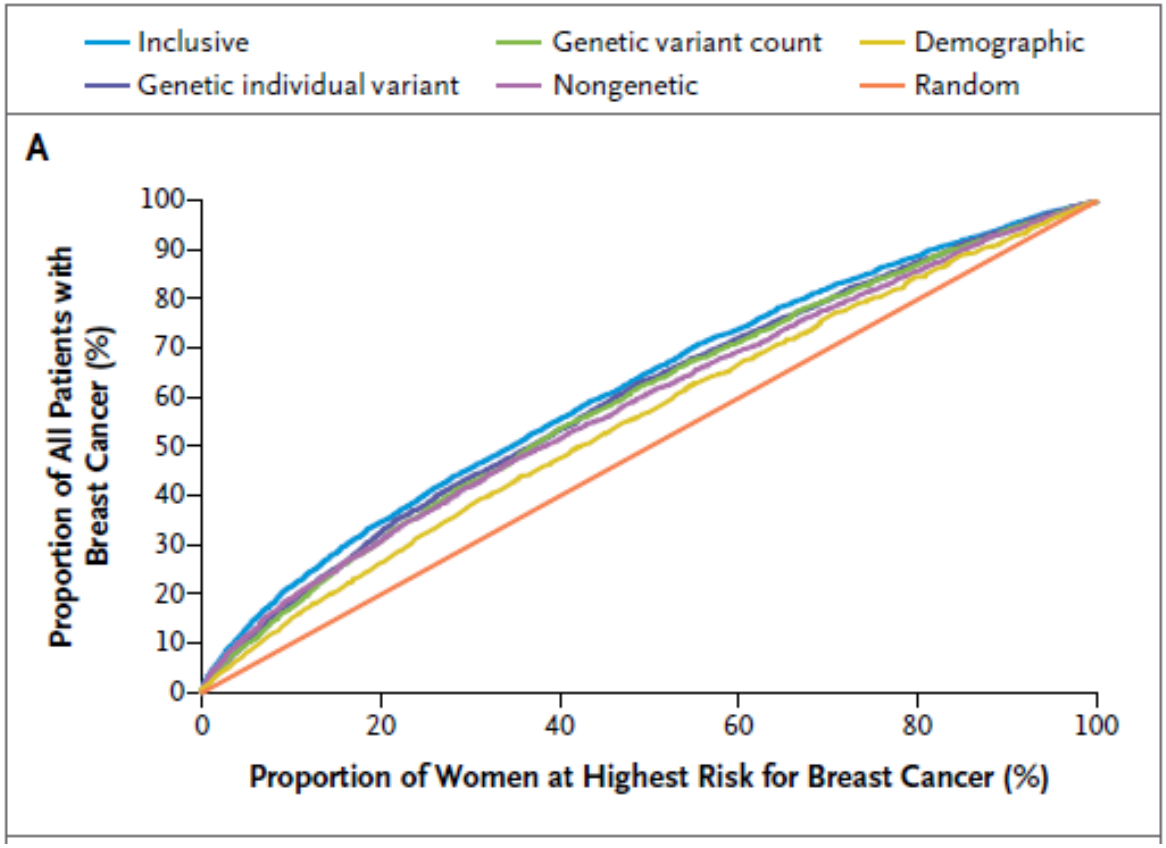
* *CASP8* denotes caspase 8, dbSNP database of single-nucleotide polymorphisms, *FGFR2* the fibroblast growth factor receptor 2 gene, *LOC643714* a hypothetical protein LOC643714, *LSP1* lymphocyte-specific protein 1, *MAP3K1* mitogen-activated protein kinase kinase kinase 1, and *TNRC9* trinucleotide repeat containing 9.

† These genes are within the linkage-disequilibrium block or blocks defined by the associated variant and are plausible candidates for the causal gene.

FIU FLORIDA INTERNATIONAL UNIVERSITY



Breast Cancer Susceptibility & GWAS



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Performance of Common Genetic Variants in Breast-Cancer Risk Models

Sholom Wacholder, Ph.D., Patricia Hartge, Sc.D., Ross Prentice, Ph.D., Montserrat Garcia-Closas, M.D., Ph.D., Heather Spencer Feigelson, Ph.D., W. Ryan Diver, M.S.P.H., Michael J. Thun, M.D., David G. Cox, Ph.D., Susan E. Hankinson, Ph.D., Peter Kraft, Ph.D., Bernard Rosner, Ph.D., Christine D. Berg, M.D., Louise A. Brinton, Ph.D., Jolanta Lissowska, Ph.D., Mark E. Sherman, M.D., Rowan Chlebowski, M.D., Charles Kooperberg, Ph.D., Rebecca D. Jackson, M.D., Dennis W. Buckman, Ph.D., Peter Hui, B.S., Ruth Pfeiffer, Ph.D., Kevin B. Jacobs, B.S., Gilles D. Thomas, M.D., Robert N. Hoover, M.D., Sc.D., Mitchell H. Gail, M.D., Ph.D., Stephen J. Chanock, M.D., and David J. Hunter, M.B., B.S., Sc.D.

FIU FLORIDA INTERNATIONAL UNIVERSITY

Breast Cancer Susceptibility & GWAS

Value of Adding Single-Nucleotide Polymorphism Genotypes to a Breast Cancer Risk Model

Mitchell H. Gail

Background Adding genotypes from seven single-nucleotide polymorphisms (SNPs), which had previously been associated with breast cancer, to the National Cancer Institute's Breast Cancer Risk Assessment Tool (BCRAT) increases the area under the receiver operating characteristic curve from 0.607 to 0.632.

Methods Criteria that are based on four clinical or public health applications were used to compare BCRAT with BCRATplus7, which includes the seven genotypes. Criteria included number of expected life-threatening events for the decision to take tamoxifen, expected decision losses (in units of the loss from giving a mammogram to a woman without detectable breast cancer) for the decision to have a mammogram, rates of risk reclassification, and number of lives saved by risk-based allocation of screening mammography. For all calculations, the following assumptions were made: Hardy–Weinberg equilibrium, linkage equilibrium across SNPs, additive effects of alleles at each locus, no interactions on the logistic scale among SNPs or with factors in BCRAT, and independence of SNPs from factors in BCRAT.

Results Improvements in expected numbers of life-threatening events were only 0.07% and 0.81% for deciding whether to take tamoxifen to prevent breast cancer for women aged 50–59 and 40–49 years, respectively. For deciding whether to recommend screening mammograms to women aged 50–54 years, the reduction in expected losses was 0.86% if the ideal breast cancer prevalence threshold for recommending mammography was that of women aged 50–54 years. Cross-classification of risks indicated that some women classified by BCRAT would have different classifications with BCRATplus7, which might be useful if BCRATplus7 was well calibrated. Improvements from BCRATplus7 were small for risk-based allocation of mammograms under costs constraints.

Conclusions The gains from BCRATplus7 are small in the applications examined. Models with SNPs, such as BCRATplus7, have not been validated for calibration in independent cohort data. Additional studies are needed to validate a model with SNPs and justify its use.

J Natl Cancer Inst 2009;101:959–963

Outline

Breast Cancer Susceptibility & GWAS

GWAS & Missing Heritability

GWAS Myths

HLA & Breast Cancer Susceptibility

GWAS & 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

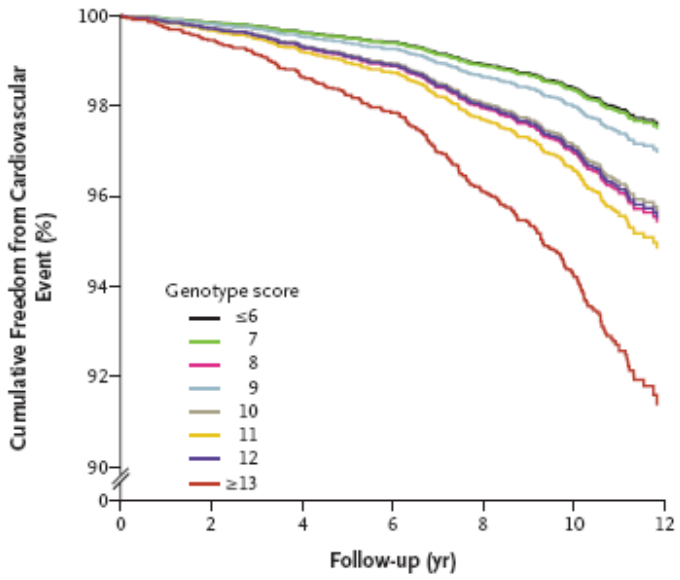
* Residual is after adjustment for age, gender, diabetes.

Finding the missing heritability of complex diseases

Teri A. Manolio[1], Francis S. Collins[2], Nancy J. Cox[3], David B. Goldstein[4], Lucia A. Hindorff[5], David J. Hunter[6], Mark I. McCarthy[7], Erin M. Ramos[5], Lon R. Cardon[8], Aravinda Chakravarti[9], Judy H. Cho[10], Alan E. Guttmacher[1], Augustine Kong[11], Leonid Kruglyak[12], Elaine Mardis[13], Charles N. Rotimi[14], Montgomery Slatkin[15], David Valle[9], Alice S. Whittemore[16], Michael Boehnke[17], Andrew G. Clark[18], Evan E. Eichler[19], Greg Gibson[20], Jonathan L. Haines[21], Trudy F. C. Mackay[22], Steven A. McCarroll[23] & Peter M. Visscher[24]

FIU | FLORIDA INTERNATIONAL UNIVERSITY

GWAS & Missing Heritability



No. at Risk

Genotype score						
≤6	122	122	120	116	112	105
7	309	304	302	298	293	267
8	574	573	566	556	538	481
9	894	894	886	872	844	757
10	913	900	887	876	849	757
11	726	719	711	696	679	604
12	465	462	456	447	434	398
≥13	229	228	218	215	207	184

Figure 1. Predicted Cumulative Freedom from Myocardial Infarction, Ischemic Stroke, or Death from Coronary Heart Disease, According to Genotype Score.

Estimates according to genotype score are derived from Cox regression models adjusted for age, sex, family history of myocardial infarction, levels of low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol, triglycerides, systolic blood pressure, diastolic blood pressure, diabetes mellitus, body-mass index, status of cigarette smoking, C-reactive protein, lipid-lowering therapy, and antihypertensive treatment.

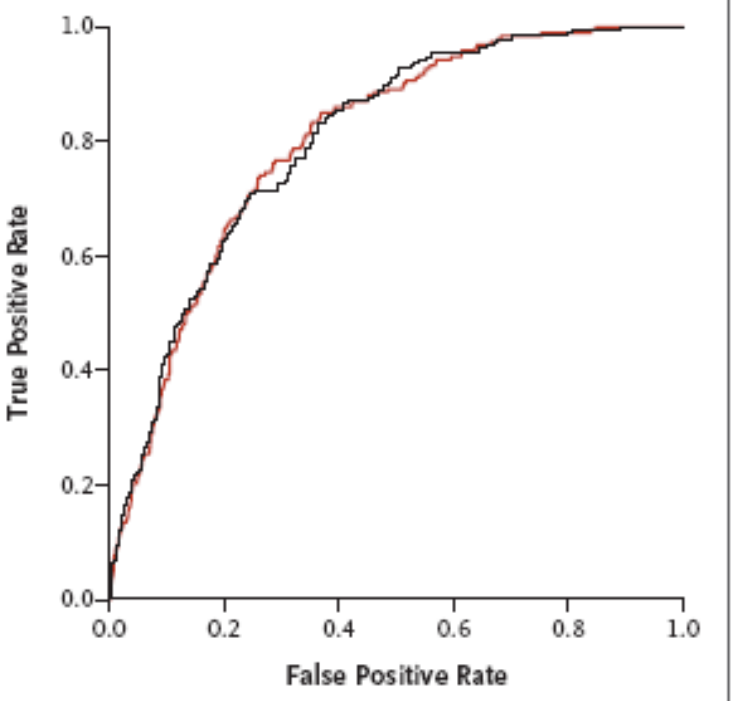


Figure 2. Receiver-Operating-Characteristic (ROC) Curves for Incident Myocardial Infarction, Ischemic Stroke, or Death from Coronary Heart Disease during 10-Year Follow-up.

The curves are based on risk-prediction models incorporating 14 clinical covariates that either included the genotype score (black line) or did not include the genotype score (red line). The C statistic (area under the ROC curve) for total cardiovascular events was the same (0.80) for both risk models.

www

Why Did GWAS Disappoint?

Table 2. Benefits, Misconceptions, and Limitations of the Genomewide Association Study.

Benefits
Does not require an initial hypothesis
Uses digital and additive data that can be mined and augmented without data degradation
Encourages the formation of collaborative consortia, which tend to continue their collaboration for subsequent analyses
Rules out specific genetic associations (e.g., by showing that no common alleles, other than *APOE*, are associated with Alzheimer's disease with a relative risk of more than 2)
Provides data on the ancestry of each subject, which assists in matching case subjects with control subjects
Provides data on both sequence and copy-number variations
Misconceptions
Thought to provide data on all genetic variability associated with disease, when in reality only common alleles with large effects are identified
Thought to screen out alleles with a small effect size, when in reality such findings may still be very useful in determining pathogenic biochemical pathways, even though low-risk alleles may be of little predictive value
Limitations
Requires samples from a large number of case subjects and control subjects and therefore can be challenging to organize
Finds loci, not genes, which can complicate the identification of pathogenic changes on an associated haplotype
Detects only alleles that are common (>5%) in a population
Requires replication in a similarly large number of samples

Why Did GWAS Disappoint?

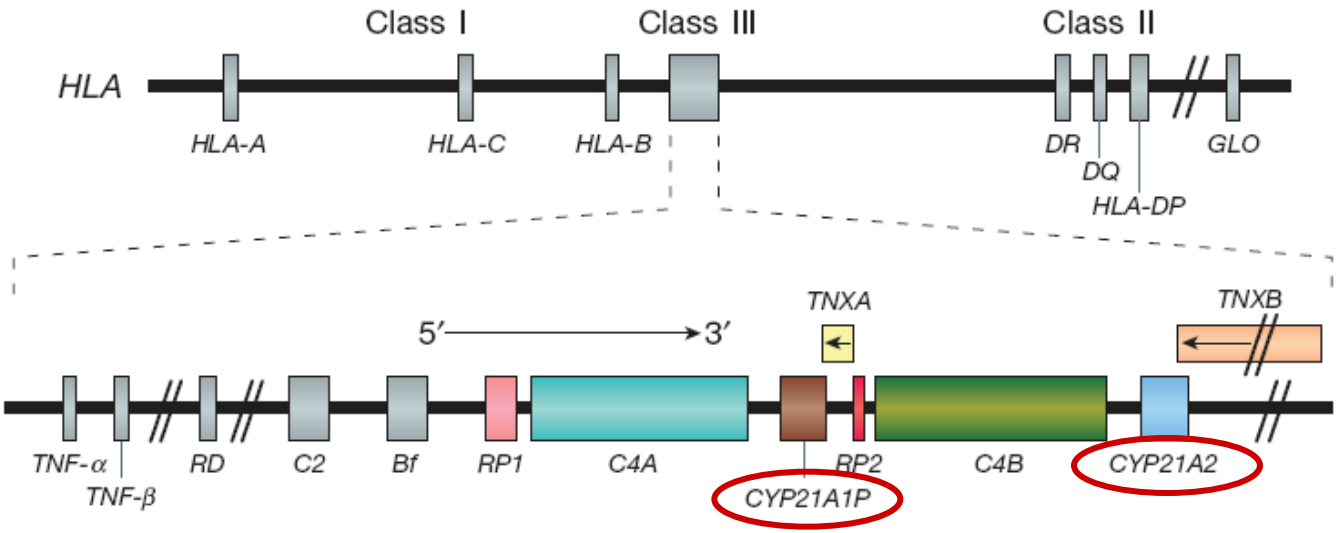
Human Molecular Genetics, 2003, Vol. 12, No. 18 **2311–2319**
DOI: 10.1093/hmg/ddg245

The HLA class III subregion is responsible for an increased breast cancer risk

Mirjam M. de Jong[1,2,3], Ilja M. Nolte[2,4], Elisabeth G. E. de Vries[1], Michael Schaapveld[6], Jan H. Kleibeuker[3], Elvira Oosterom[4], Jan C. Oosterwijk[2], Annemarie H. van der Hout[2], Gerrit van der Steege[4], Marcel Bruinenberg[4], H. Marike Boezen[5], Gerard J. te Meerman[2] and Winette T. A. van der Graaf[1,*]

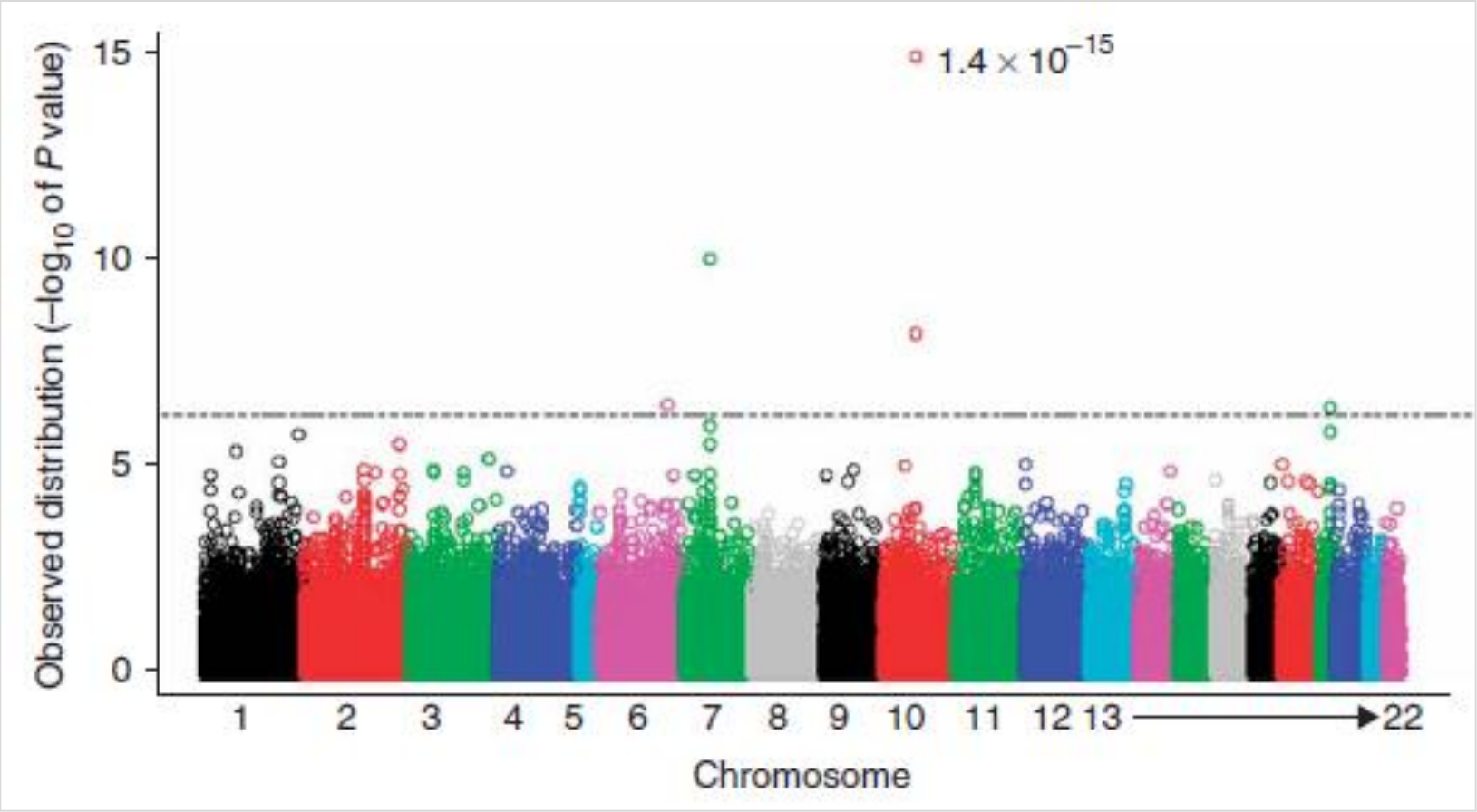
[1]Department of Medical Oncology, [2]Department of Medical Genetics, [3]Department of Gastroenterology, [4]Department of Medical Biology and [5]Department of Epidemiology, University Medical Center, Groningen, The Netherlands and [6]Comprehensive Cancer Center Northern Netherlands, Groningen, The Netherlands

Our strongest candidate gene CYP21A2 is in the class III region….



but not included in GWAS
because of the presence of a paralog in the genome

GWAS in Childhood Leukemia



FIU FLORIDA INTERNATIONAL UNIVERSITY

Germline genomic variants associated with childhood acute lymphoblastic leukemia

Lisa R Treviño[1,6], Wenjian Yang[1,6], Deborah French[1], Stephen P Hunger[2], William L Carroll[3], Meenakshi Devidas[4], Cheryl Willman[5], Geoffrey Neale[1], James Downing[1], Susana C Raimondi[1], Ching-Hon Pui[1], William E Evans[1] & Mary V Relling[1]

GWAS Analysis

Common SNPs explain a large proportion of the heritability for human height

Jian Yang[1], Beben Benyamin[1], Brian P McEvoy[1], Scott Gordon[1], Anjali K Henders[1], Dale R Nyholt[1], Pamela A Madden[2], Andrew C Heath[2], Nicholas G Martin[1], Grant W Montgomery[1], Michael E Goddard[3] & Peter M Visscher[1]

Highly significant and well-replicated SNPs identified to date explain only ~5% of the phenotypic variance for height[19]. Our results show that common SNPs in total explain another ~40% of phenotypic variance. Hence, 88% (40/45) of the variation due to SNPs has been undetected in published GWASs because the effects of the SNPs are too small to be statistically significant. Our results also suggest that the discrepancy between 80% heritability and 45% accounted for by all SNPs is due to incomplete LD between causal variants and the SNPs, possibly because the causal variants have a lower MAF on average than the SNPs typed on the array. We cannot tell from these results whether or not some of this discrepancy is due to causal variants with very low frequency—for example, MAF < 0.001 (ref. 4).

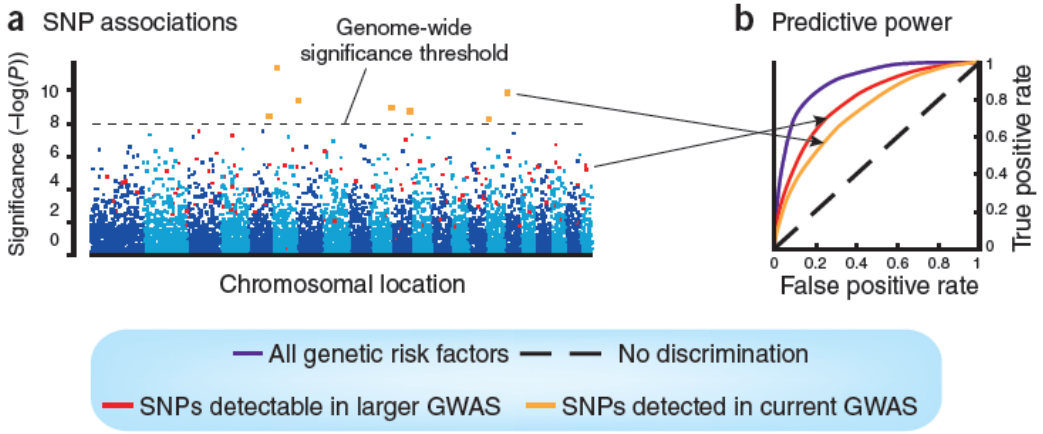


Figure 1 Predictive power of genetic variants identified through genome-wide association studies. (**a**) GWAS typically report a visualization of genome-wide association of SNPs using a Manhattan plot of significance against chromosomal location, shown here for a generic common disease. Gold, SNPs detected in current GWAS; red, SNPs that may be detectable in GWAS with larger sample sizes. (**b**) SNPs identified through GWAS as significantly associated with disease susceptibility may be used in a genetic predictive test to classify disease risk in individual. Using SNPs identified from current GWAS increases the area under the curve (AUC), reflecting increased ability to classify risk (gold line) over nondiscrimination (dashed black line). Discovery of additional associated SNPs through larger GWAS (red curve) further increases the AUC, improving the genetic test, but does not approach what would be theoretically possible with full characterization of all genetic variation influencing the disease (purple).

FIU | FLORIDA INTERNATIONAL UNIVERSITY

GWAS Analysis

OPEN ACCESS Freely available online

PLoS one

Estimating the Total Number of Susceptibility Variants Underlying Complex Diseases from Genome-Wide Association Studies

Hon-Cheong So[1], Benjamin H. K. Yip[1], Pak Chung Sham[1,2,3]*

1 Department of Psychiatry, University of Hong Kong, Hong Kong SAR, China, 2 Genome Research Centre, University of Hong Kong, Hong Kong SAR, China, 3 The State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong SAR, China

Abstract

Recently genome-wide association studies (GWAS) have identified numerous susceptibility variants for complex diseases. In this study we proposed several approaches to estimate the total number of variants underlying these diseases. We assume that the variance explained by genetic markers (Vg) follow an exponential distribution, which is justified by previous studies on theories of adaptation. Our aim is to fit the observed distribution of Vg from GWAS to its theoretical distribution. The number of variants is obtained by the heritability divided by the estimated mean of the exponential distribution. In practice, due to limited sample sizes, there is insufficient power to detect variants with small effects. Therefore the power was taken into account in fitting. Besides considering the most significant variants, we also tried to relax the significance threshold, allowing more markers to be fitted. The effects of false positive variants were removed by considering the local false discovery rates. In addition, we developed an alternative approach by directly fitting the *z*-statistics from GWAS to its theoretical distribution. In all cases, the "winner's curse" effect was corrected analytically. Confidence intervals were also derived. Simulations were performed to compare and verify the performance of different estimators (which incorporates various means of winner's curse correction) and the coverage of the proposed analytic confidence intervals. Our methodology only requires summary statistics and is able to handle both binary and continuous traits. Finally we applied the methods to a few real disease examples (lipid traits, type 2 diabetes and Crohn's disease) and estimated that hundreds to nearly a thousand variants underlie these traits.

Table 7. Estimated number of susceptibility variants assuming a gamma distribution of effect sizes.

	Shape	Lambda	Mean	Number of variants
LDL	0.9	937	9.60E-04	375
	0.7	845	8.28E-04	435
	0.5	754	6.63E-04	543
	0.3	664	4.52E-04	797
HDL	0.9	657	1.37E-03	460
	0.7	585	1.20E-03	526
	0.5	514	9.73E-04	648
	0.3	445	6.74E-04	935
TG	0.9	1152	7.81E-04	474
	0.7	1038	6.75E-04	548
	0.5	812	6.16E-04	601
	0.3	924	3.25E-04	1140
Crohn (all)	0.9	1328	6.78E-04	812
	0.7	1211	5.78E-04	951
	0.5	1093	4.57E-04	1203
	0.3	970	3.09E-04	1779
Crohn (pruned)	0.9	1649	5.46E-04	1008
	0.7	1512	4.63E-04	1188
	0.5	1375	3.64E-04	1513
	0.3	1219	2.46E-04	2235
DM (all)	0.9	1122	8.02E-04	528
	0.7	1042	6.72E-04	631
	0.5	962	5.20E-04	816
	0.3	884	3.40E-04	1249
DM (pruned)	0.9	1165	7.73E-04	549
	0.7	1095	6.39E-04	663
	0.5	1026	4.87E-04	870
	0.3	994	3.02E-04	1405

When the shape parameter equals one, the gamma distribution is equivalent to an exponential distribution and the results are listed in table 6. When the shape parameter decreases, the distribution is more skewed towards zero, implying that we assume more variants to have small effect sizes.

FIU FLORIDA INTERNATIONAL UNIVERSITY

Outline

Breast Cancer Susceptibility & GWAS

GWAS & Missing Heritability

GWAS Myths

HLA & Breast Cancer Susceptibility

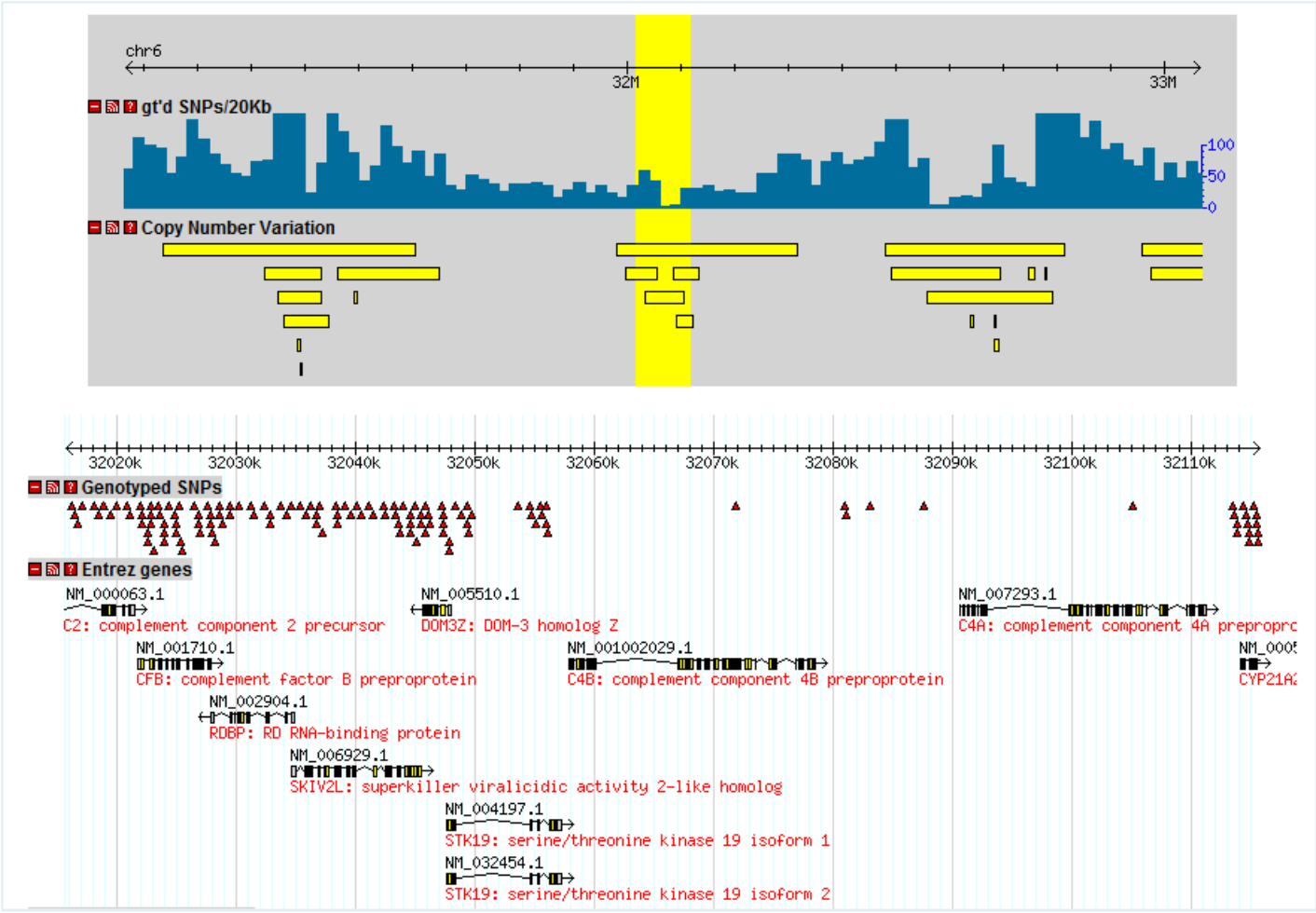
GWAS Myths

Genome-wide coverage of all genes and variants

No need for candidate gene studies

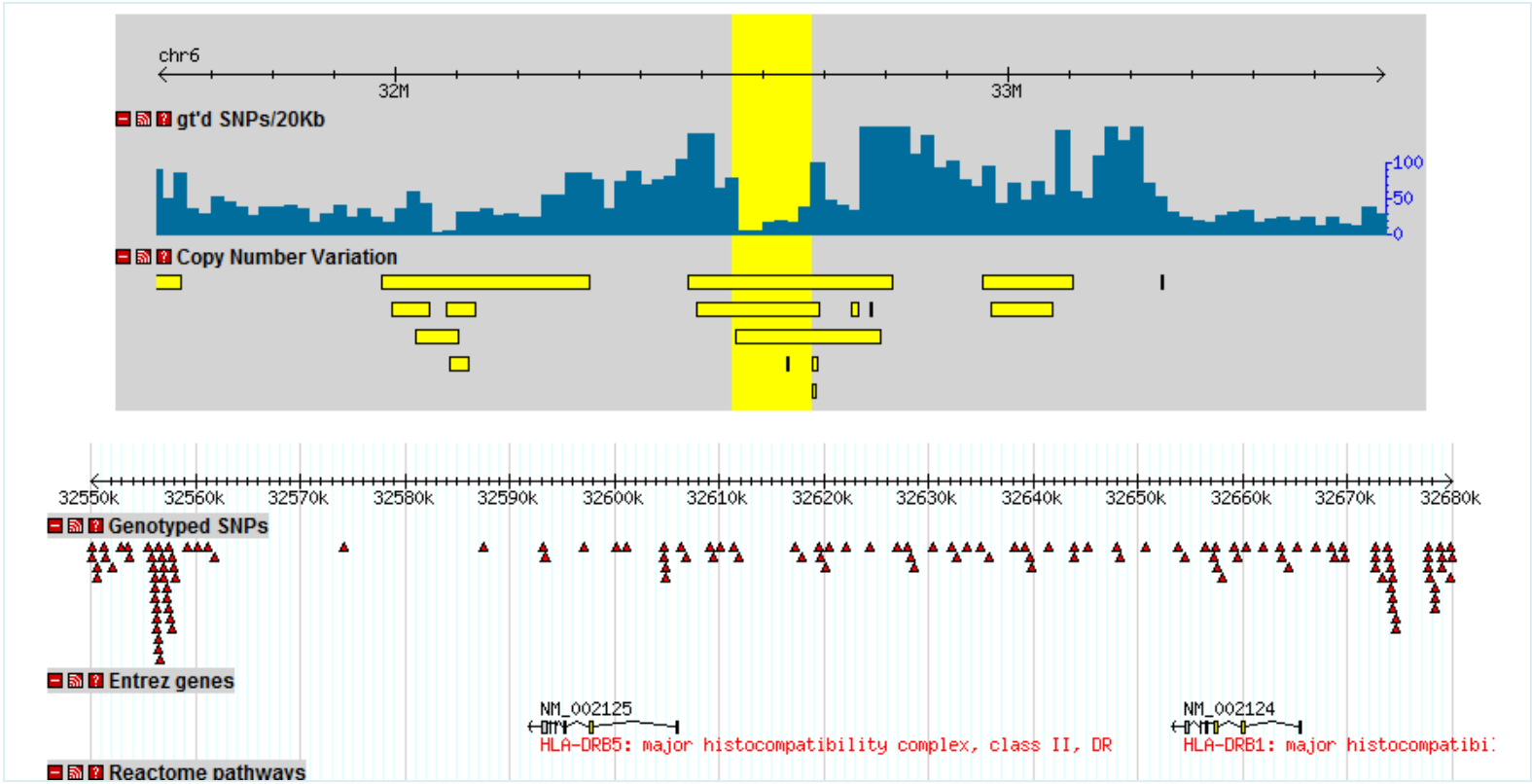
If GWAS do not find an association, no other study will find

What does GWAS Offer for HLA Associations



FIU | FLORIDA INTERNATIONAL UNIVERSITY

What does GWAS Offer for HLA Associations



FIU | FLORIDA INTERNATIONAL UNIVERSITY

OPEN ACCESS Freely available online

PLoS one

High Frequency of Copy Number Variations and Sequence Variants at *CYP21A2* Locus: Implication for the Genetic Diagnosis of 21-Hydroxylase Deficiency

Silvia Parajes[1], Celsa Quinteiro[1], Fernando Domínguez[1,2], Lourdes Loidi[1]*

1 Fundación Pública Galega de Medicina Xenómica (Unidad de Medicina Molecular), Hospital Clínico Universitario, Santiago de Compostela, A Coruña, Spain, **2** Departamento de Fisiología, Universidad de Santiago de Compostela, Santiago de Compostela, A Coruña, Spain

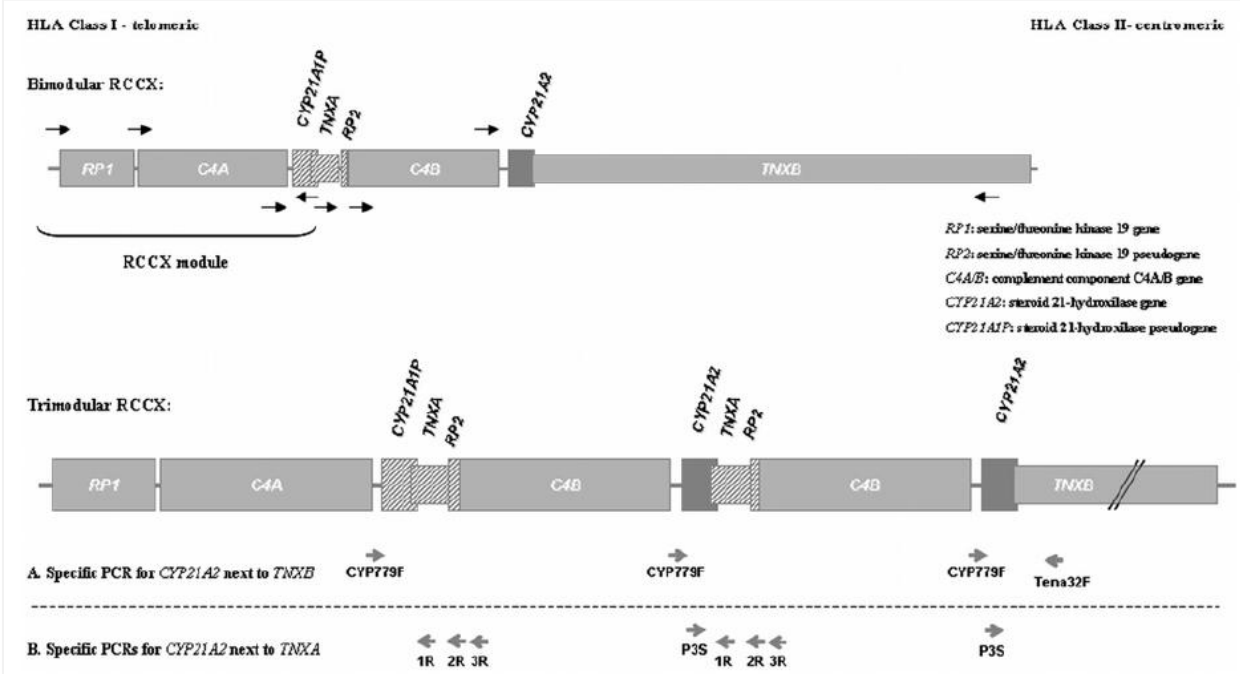


Figure 1. Organization of RCCX module and Copy Number Variations at chromosome 6p21. Black horizontal arrows denote gene orientation. A. Specific primers used for the amplification of *CYP21A2* next to *TNXB*. B. Specific primers used for the amplification of *CYP21A2* next to *TNXA*. Primer binding sites for each primer are indicated by grey horizontal arrows.
doi:10.1371/journal.pone.0002138.g001

FIU FLORIDA INTERNATIONAL UNIVERSITY

Outline

Breast Cancer Susceptibility & GWAS

GWAS & Missing Heritability

GWAS Myths

HLA & Breast Cancer Susceptibility

HLA & Breast Cancer Susceptibility

What is the HLA complex and why is almost any disease risk associated with it?

HLA & Breast Cancer Susceptibility

Relevant non-HLA genes

Why is There an HLA Association in Almost Any Disease?

The very first MHC association was with leukemia in mice (1964) and with Hodgkin disease in humans (1967)

Many cancers show associations and some (NPC) even show linkage to MHC

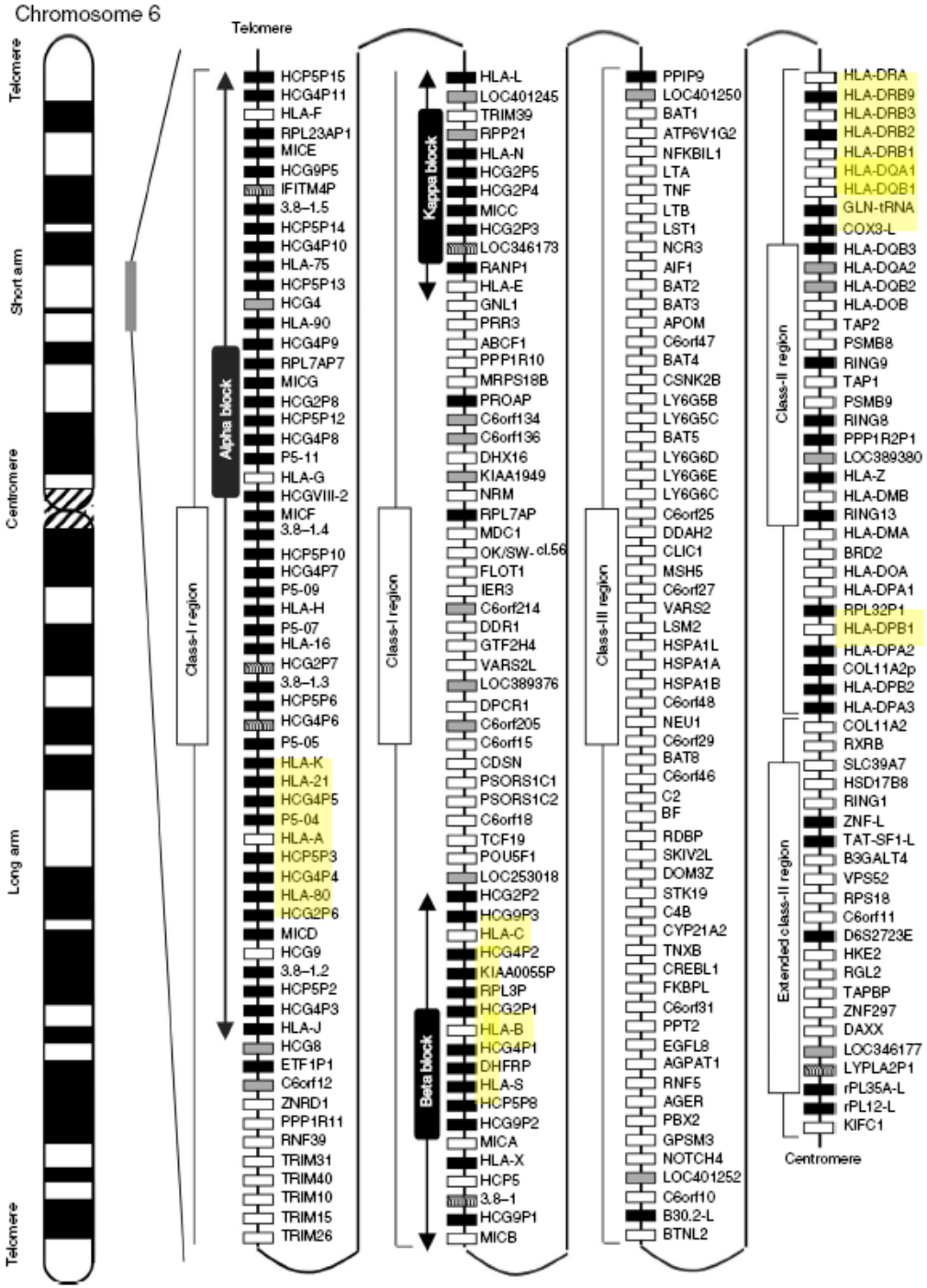
CAH and HH genes were first located in and around HLA by association studies

Autoimmune disorders and infectious diseases show the strongest associations

Besides, sarcoidosis, birth weight, obesity, long QT syndrome and many others show associations with HLA alleles or haplotypes

(in the 1980s, we did not know much about non-HLA genes)

But now we do!



Shiina et al, 2004 (www)

Transcription Factors in the Extended MHC

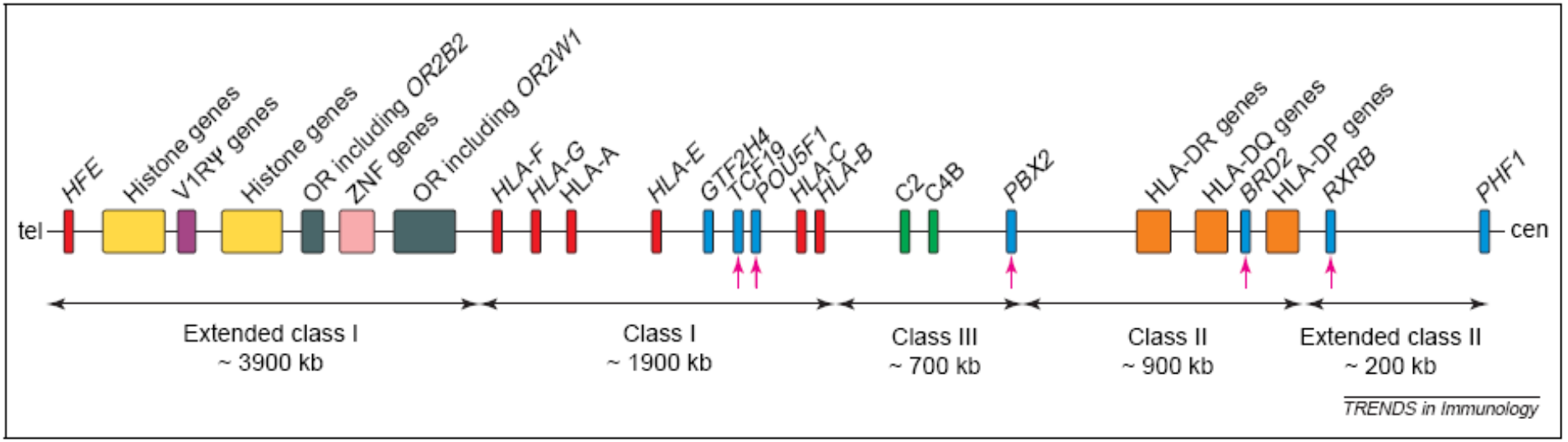


Figure 1. Map of the extended human MHC. The map (not to scale) shows selected genes and gene clusters of the extended MHC (xMHC) from telomere (tel, left) to centromere (cen, right) on the short arm of human chromosome 6. The total number of genes encoded within the xMHC is 578 [26]. The five subregions making up the xMHC span ~7600 kilobasepairs (Kb) and are indicated by arrows below the map, with their approximate lengths. The following types of genes are mentioned within the review: class I genes (red), class II genes (orange), OR gene clusters (dark green), V1R pseudogene cluster (violet), zinc finger genes (pink, only one of the several locations of ZNF loci is shown) and TF genes (blue). The red arrows indicate those TF genes whose location within the xMHC is conserved evolutionarily from fish to mammals. The following genes with their symbols are depicted: *HFE*, hemochromatosis; *OR2B2*, olfactory receptor, family 2, subfamily B, member 2; *OR2W1*, olfactory receptor, family 2, subfamily W, member 1; *GTF2H4*, general transcription factor IIH, polypeptide 4; *POU5F1*, Pou domain, class 5, transcription factor 1; *TCF19*, transcription factor 19; *C2*, complement component 2; *C4B*, complement component 4B; *PBX2*, pre-B-cell leukemia transcription factor 2; *BRD2*, bromodomain-containing protein 2; *RXRB*, retinoid X receptor, β; *PHF1*, PHD finger protein 1. The *POU5F1* gene is also known as *Oct4* in the mouse. Further details can be found in recently published reviews [26,55].

Ziegler, 2005 (www)

HLA Polymorphisms

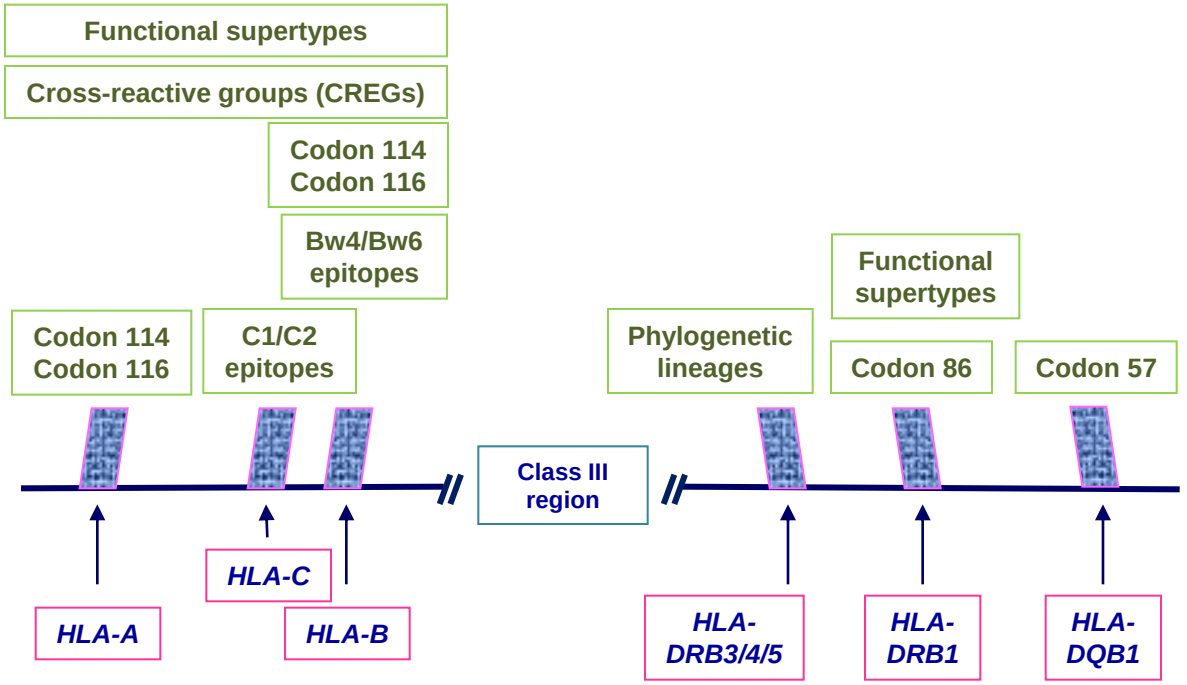
- Highest resolution DNA level HLA alleles: Related to transplantation success, susceptibility to diseases related to antigen presentation (autoimmune disorders, infectious diseases)
- Serological HLA antigens: Relevant to transplantation and disease associations
- HLA epitopes (Bw4/Bw6; C1/C2): Interactions with NK cell
- Functional supertypes: Involved in antigen presentation
- Genetic supertypes: Represents ancestral lineages

HLA Polymorphisms

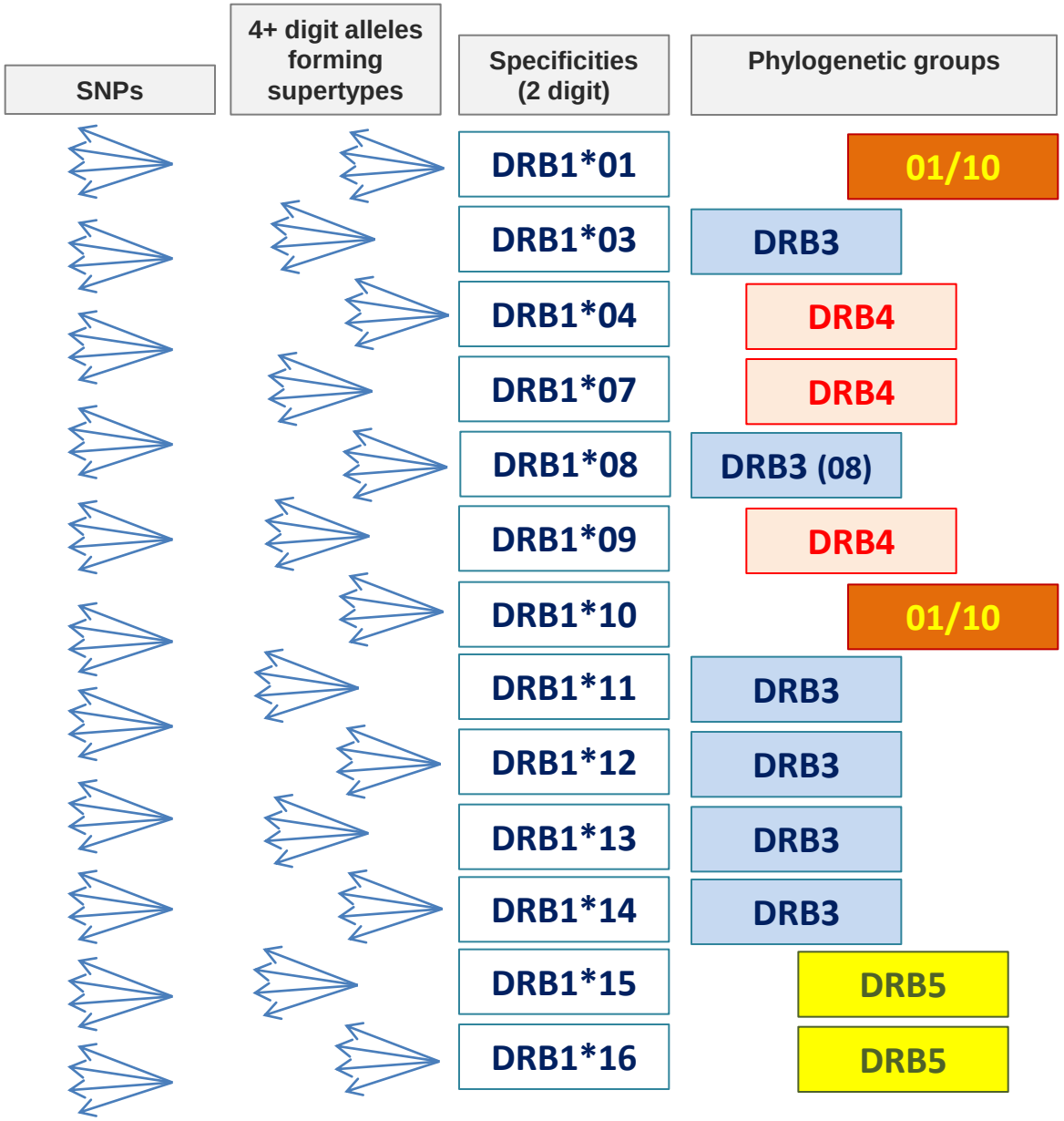
Current disease association studies are mainly concerned with high resolution allelic associations and may miss out a lot of information

Examination of functional groupings and lineages rather than individual alleles may be a more powerful approach

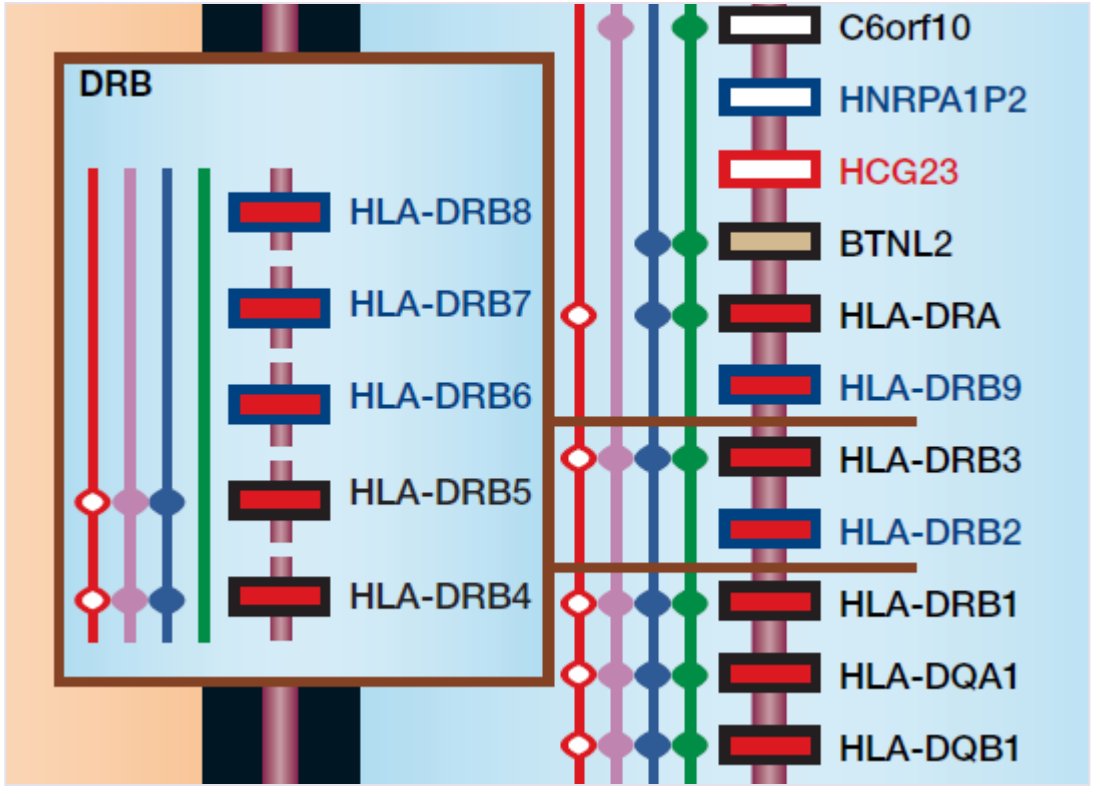
Functional Multi-allelic HLA Polymorphisms



Functional Multi-allelic HLA Polymorphisms



HLA-DR/DQ Region



FIU FLORIDA INTERNATIONAL UNIVERSITY

nature REVIEWS GENETICS

Gene map of the extended human MHC

Roger Horton, Laurens Wilming, Vikki Rand, Ruth C. Lovering, Elspeth A. Bruford, Varsha K. Khodiyar, Michael J. Lush, Sue Povey, C. Conover Talbot Jr. , Mathew W. Wright, Hester M. Wain, John Trowsdale, Andreas Ziegler and Stephan Beck.

The extended major histocompatibility complex (xMHC) on the short arm of human chromosome 6 is essential for adaptive and innate immunity. In addition to their vital role in transplant medicine, certain combinations (haplotypes) of xMHC loci are known to confer protection from, or susceptibility to, many diseases including most if not all, autoimmune, inflammatory and infectious diseases. Paralleled only by war and famine, these diseases rank as the leading cause of human mortality and disability worldwide. History has shown that knowledge of detailed and accurate genetic maps enhances our ability to diagnose, understand and treat disease at the molecular level.

The genetic map presented here aims to be the definitive protein-coding gene map of the human xMHC. It comprises 421 loci, of which 252 are classified as expressed genes, 139 as pseudogenes and 30 as transcripts (based on EST evidence, but without open reading frames). This represents an increase of 345 annotated loci since the first MHC poster in 1991 (REF. 1) and juxtaposes the classical MHC with several gene clusters, including some of the largest in the human genome. The accompanying article² reviews the current state of knowledge relating to these loci in the context of MHC biology and disease.

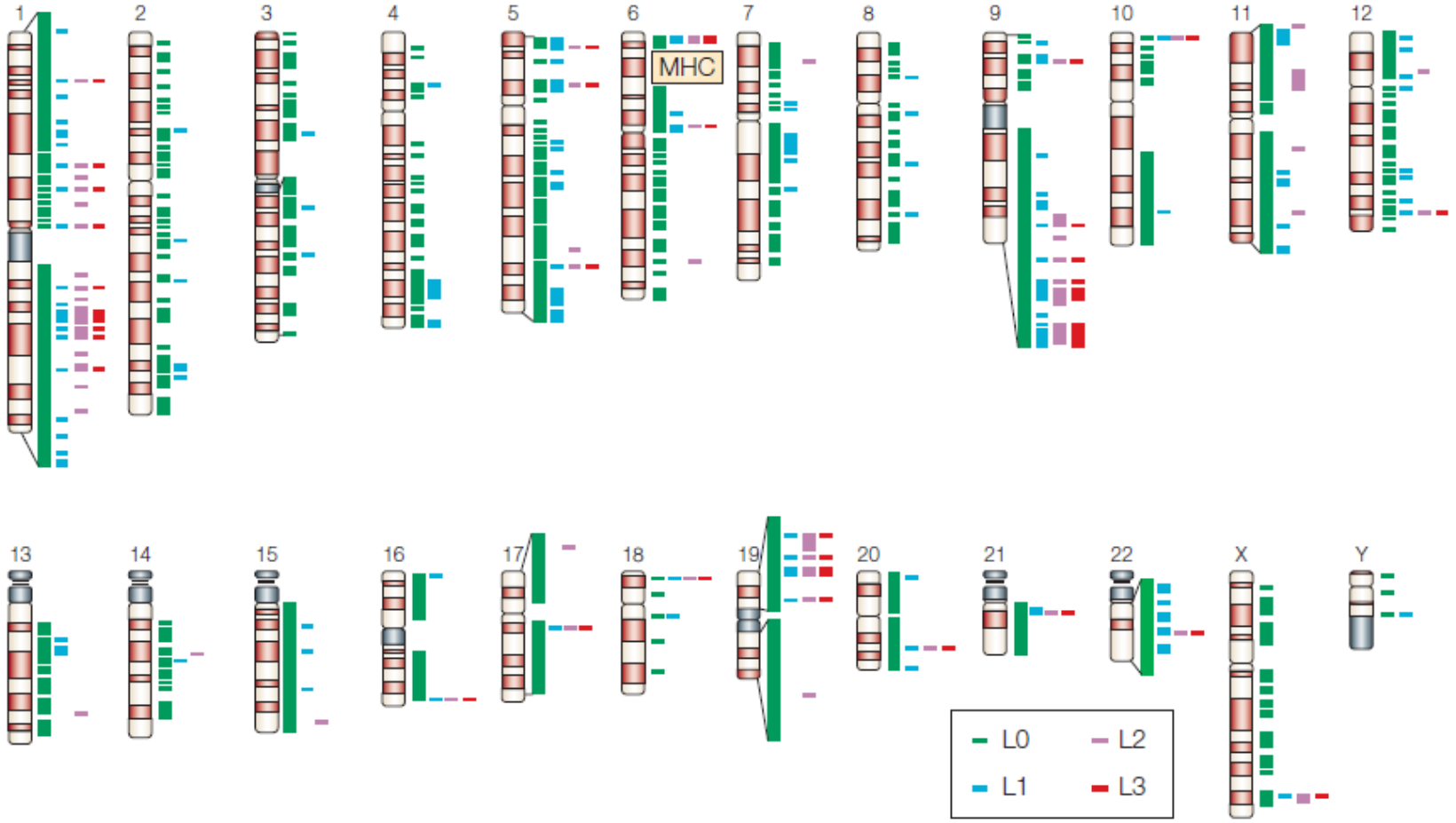


Figure 2 | **Distribution of major histocompatibility complex (MHC) paralogues in the human genome.** The approximate positions of the putative paralogues are colour-coded according to confidence level: L0 column represents BLAST similarity matches with a *p*-value of less than 10^{-5} (green); L1 column represents BLAST matches after filtering out domain-only matches (blue); L2 column represents BLAST matches after filtering for conserved gene structure[100] (purple); L3 column represents BLAST matches that passed both filtering steps (red).

nature REVIEWS GENETICS

Gene map of the extended human MHC

Roger Horton, Laurens Wilming, Vikki Rand, Ruth C. Lovering, Elspeth A. Bruford, Varsha K. Khodiyar, Michael J. Lush, Sue Povey, C. Conover Talbot Jr. , Mathew W. Wright, Hester M. Wain, John Trowsdale, Andreas Ziegler and Stephan Beck.

The extended major histocompatibility complex (xMHC) on the short arm of human chromosome 6 is essential for adaptive and innate immunity. In addition to their vital role in transplant medicine, certain combinations (haplotypes) of xMHC loci are known to confer protection from, or susceptibility to, many diseases including most if not all, autoimmune, inflammatory and infectious diseases. Paralleled only by war and famine, these diseases rank as the leading cause of human mortality and disability worldwide. History has shown that knowledge of detailed and accurate genetic maps enhances our ability to diagnose, understand and treat disease at the molecular level.

The genetic map presented here aims to be the definitive protein-coding gene map of the human xMHC. It comprises 421 loci, of which 252 are classified as expressed genes, 139 as pseudogenes and 30 as transcripts (based on EST evidence, but without open reading frames). This represents an increase of 345 annotated loci since the first MHC poster in 1991 (REF. 1) and juxtaposes the classical MHC with several gene clusters, including some of the largest in the human genome. The accompanying article[2] reviews the current state of knowledge relating to these loci in the context of MHC biology and disease.

Human Major Histocompatibility Complex

Most gene-dense region in the genome

Table 1. Human Genome Top 10 Gene-Dense Regions

GoldenPath location	Region	%GC	% repeats	Genes/Mb	Comments
chr6:31250001–32500000	HLAC–HLADRB3	47	47	48.8	Includes MHC class III region
chr6:25500001–26500000	FLJ20048–BTN2A3	41	43	44.0	Includes histone families
chr12:6250001–7250000	FLJ10665–PXR1	46	41	43.1	Includes CD4, complement 1
chr17:39000001–40000000	KRT23–ACLY	46	44	43.0	Includes keratin families
chr19:53250001–55000000	ELSPBP1–TCBAP0758	52	57	42.3	Includes CD37
chr16:250001–1500000	DKFZP761D0211–KIAA0683	60	28	40.8	GC rich
chr11:250001–1500000	AP2A2–HCCA2	53	36	40.2	Gap in sequence; includes IRF7, TOLLIP
chr17:7000001–8000000	ASGR1–PER1	51	43	39.0	Includes TNSF12, 13; CD68; TP53
chrX:150500001–151500000	DUSP9–GAB3	53	43	39.0	Includes G6PD; IRAK1
chr19:59250001–60250000	OSCAR–RDH13	49	53	36.0	Includes KIR, ILT, LILR families

Using a window offset of 250 kb, the number of genes per megabase and GC content were calculated as described in Figure 1. If a region appeared in the top 20 hits more than once (e.g., chr16:250001–250000 and chr16:5000001–1,500000), the regions were combined. "Region" indicates the outermost genes within the GoldenPath span.

Xie, 2003 (www)

HLA & Breast Cancer Susceptibility

Genetic susceptibility to breast cancer: HLA DQB*03032 and HLA DRB1*11 may represent protective alleles

Subhra Chaudhuri, Annaiah Cariappa, Mei Tang, Daphne Bell, Daniel A. Haber, Kurt J. Isselbacher, Dianne Finkelstein, David Forcione, and Shiv Pillai†

Massachusetts General Hospital Cancer Center and Harvard Medical School, Building 149, 13th Street, Charlestown, MA 02129

PNAS | October 10, 2000 | vol. 97 | no. 21 | 11451–11454

Human Molecular Genetics, 2003, Vol. 12, No. 18 **2311–2319**
DOI: 10.1093/hmg/ddg245

The HLA class III subregion is responsible for an increased breast cancer risk

Mirjam M. de Jong[1,2,3], Ilja M. Nolte[2,4], Elisabeth G. E. de Vries[1], Michael Schaapveld[6], Jan H. Kleibeuker[3], Elvira Oosterom[4], Jan C. Oosterwijk[2], Annemarie H. van der Hout[2], Gerrit van der Steege[4], Marcel Bruinenberg[4], H. Marike Boezen[5], Gerard J. te Meerman[2] and Winette T. A. van der Graaf[1,*]

[1]Department of Medical Oncology, [2]Department of Medical Genetics, [3]Department of Gastroenterology, [4]Department of Medical Biology and [5]Department of Epidemiology, University Medical Center, Groningen, The Netherlands and [6]Comprehensive Cancer Center Northern Netherlands, Groningen, The Netherlands

Non-HLA Genes of the HLA Complex Involved in Fundamental Cellular Processes

- **transcriptional or translational machinery** (*GTF2H4, TCF19, POU5F1, ZNRD1, LSM2, BAT1, RDBP, VARS, PBX2, DOM3Z, SKIV2L, DHX16, GNL1, RPS18, MRPS18B; CSNK2B, TRIM26 , BRD2, PHF1, CREBL1, BTK19, RXRB, STK19, ABF1)*
- **house-keeping** (*DOM3Z, NEU1, AGPAT1, CLIC1, CSNK2B*)
- **biosynthesis, electron transport and hydrolase activity** (*PPT2, DDAH2, ATP6V1G2*)
- **protein–protein interactions, chaperone function, ubiquitination and signalling** (*ZBTB12 (C6orf46), HSPA1A, HSPA1B, BAT3, BAT8, AGAR, RNF5, FKBPL, LST1, TNXB* and *NOTCH4*)
- **genome surveillance machinery and chromosome stability** (*MDC1, MSH5, GTF2H4; DAXX; UBD; -CDKN1A-*)
- **apoptosis** (*BAT2, BAT3, LTA/LTB, IER3, DAXX, DDR1; -CDKN1A-*)
- **cell cycle regulation** (*TCF19; ZNRD1; CSNK2B; CLIC1; FKRPL; -CDKN1A-)*
- **cell division** (*KIFC1*)
- **meiosis** (*MSH5*)
- **spermatogenesis or sperm motility** (*SKIV2L; CLIC1; HSPA1B; -TCP11-)*
- **embryonic expression** (*DAXX, HSPA1A/B; NOTCH4*)
- **multidrug resistance** (*ZNRD1, MSH5, TAP1, TAP2*)
- **angiogenesis** (*NOTCH4, -EDN1-*)
- **proto-oncogenes** (*NOTCH4, PBX2*)
- **hormonal effects** (*CYP21A2, HSD17B8*)
- **immunoregulatory role** (*C2, C4, CFB, LTA; TNF; LTB; CLIC1; IER3; MYLIP; UBD; FKBPL; TAP1, TAP2, TAPBP, PSMB8, PSMB9, NEU1; PRSS16; HLA-E; HLA-DMA, HLA-DMB, HLA-DOA, HLA-DOB, -CDKN1A-*)
- **inflammation** (*LTA; TNF; LTB; AIF1; NFKBIL1; BAT1; DDAH2; CLIC1; ABCF1*)
- **radioresistance** (*FKBPL; MDC1*)

Table 1: Candidate genes in the extended HLA complex	
NOTCH4	Human homolog of the MMTV integration site *Int3*, and involved in breast development as well as angiogenesis[7]
MDC1	Major DNA repair pathway protein[8] whose overexpression is a BC prognostic marker[9]. *MSH5* is another DNA repair gene in the HLA complex.
DDR1	Discoidin domain receptor 1; CD167; also known as mammary carcinoma kinase 10 (MCK10) [10]
UBD, IER3, BAT3, AIF1, RXRB, DAXX	HLA region genes participating in apoptosis, cell differentiation some through interaction with TP53 pathway[11,12]
HSPA1B	Encodes HSP70; a variant is associated with BC risk and not in GWAS chips; expression predicts prognosis[13]
EHMT2	Much more active one of the two histone-lysine N-methyltransferases. A variant is associated with BC susceptibility[14]
CYP21A2	21-hydroxylase; the most important enzyme involved in adrenal sex steroid biosynthesis
HSD17B8	Sex steroid interconversion
HFE	Its mutations induce iron overload and are associated with BC susceptibility[15-17]
SLC39A4	Mediates increased growth factor receptor activation via zinc-induced inhibition of phosphatases, leading to increased growth and invasion in breast cancer cells in vitro[18]

FEBS 17840 FEBS Letters 398 (1996) 165–169

Receptor protein tyrosine kinase *DDR* is up-regulated by p53 protein

Shirou Sakuma[a,*], Hideyuki Saya[b], Mitsuhiro Tada[a], Mitsuyoshi Nakao[b], Toshiyoshi Fujiwara[c], Jack A. Roth[d], Yutaka Sawamura[a], Yumiko Shinohe[a], Hiroshi Abe[a]

[a] *Department of Neurosurgery, University of Hokkaido School of Medicine, North 15, West 7, Kita-ku, Sapporo 060, Japan*
[b] *Department of Tumor Genetics and Biology, Kumamoto University School of Medicine, Kumamoto 860, Japan*
[c] *First Department of Surgery, Okayama University Medical School, Okayama 700, Japan*
[d] *Section of Thoracic Molecular Oncology, Department of Thoracic and Cardiovascular Surgery, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030 USA*

The EMBO Journal Vol. 22 No. 6 pp. 1289–1301, 2003

p53 induction and activation of DDR1 kinase counteract p53-mediated apoptosis and influence p53 regulation through a positive feedback loop

Pat P.Ongusaha, Jong-il Kim[1], Li Fang[2], Tai W.Wong[3], George D.Yancopoulos[4], Stuart A.Aaronson[2] and Sam W.Lee[5]

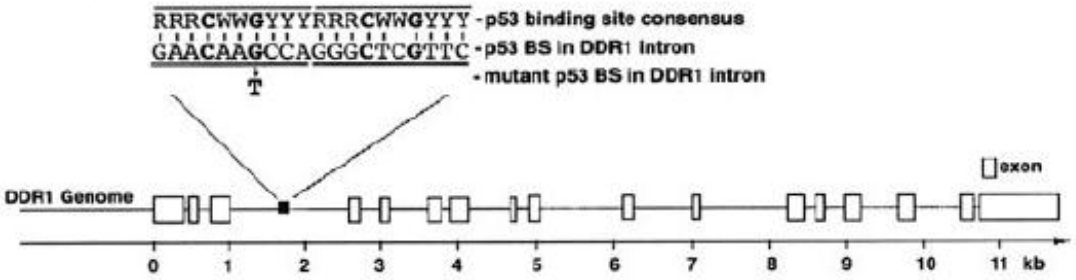
Cancer Biology Program, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine and Harvard Medical School, Boston, MA 02115, [2]Derald H. Ruttenberg Cancer Center, Mount Sinai School of Medicine, New York, NY 10029, [3]Oncology Drug Discovery Group, Bristol-Meyer Squibb Pharmaceutical Research Institutes, Princeton, NJ 08543 and [4]Regeneron Pharmaceuticals, Inc., Tarrytown, NY 10591, USA

MOLECULAR AND CELLULAR BIOLOGY, Apr. 2001, p. 2906–2917 Vol. 21, No. 8
0270-7306/01/$04.00+0 DOI: 10.1128/MCB.21.8.2906–2917.2001

Discoidin Domain Receptor 1 Tyrosine Kinase Has an Essential Role in Mammary Gland Development

WOLFGANG F. VOGEL,[1]* ATTILA ASZÓDI,[2] FRAUKE ALVES,[3] AND TONY PAWSON[1,4]

Programme in Molecular Biology and Cancer, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1X5,[1] and Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5S 1A8,[4] Canada; Department of Experimental Pathology, Lund University Hospital, 22185 Lund, Sweden[2]; and Department of Hematology and Oncology, University of Göttingen, 37075 Göttingen, Germany[3]



MHC & DNA Repair

1: Tissue Antigens. 1981 Jan;17(1):104-10.

DNA repair, H-2, and aging in NZB and CBA mice.

Hall KY, Bergman K, Walford RL.

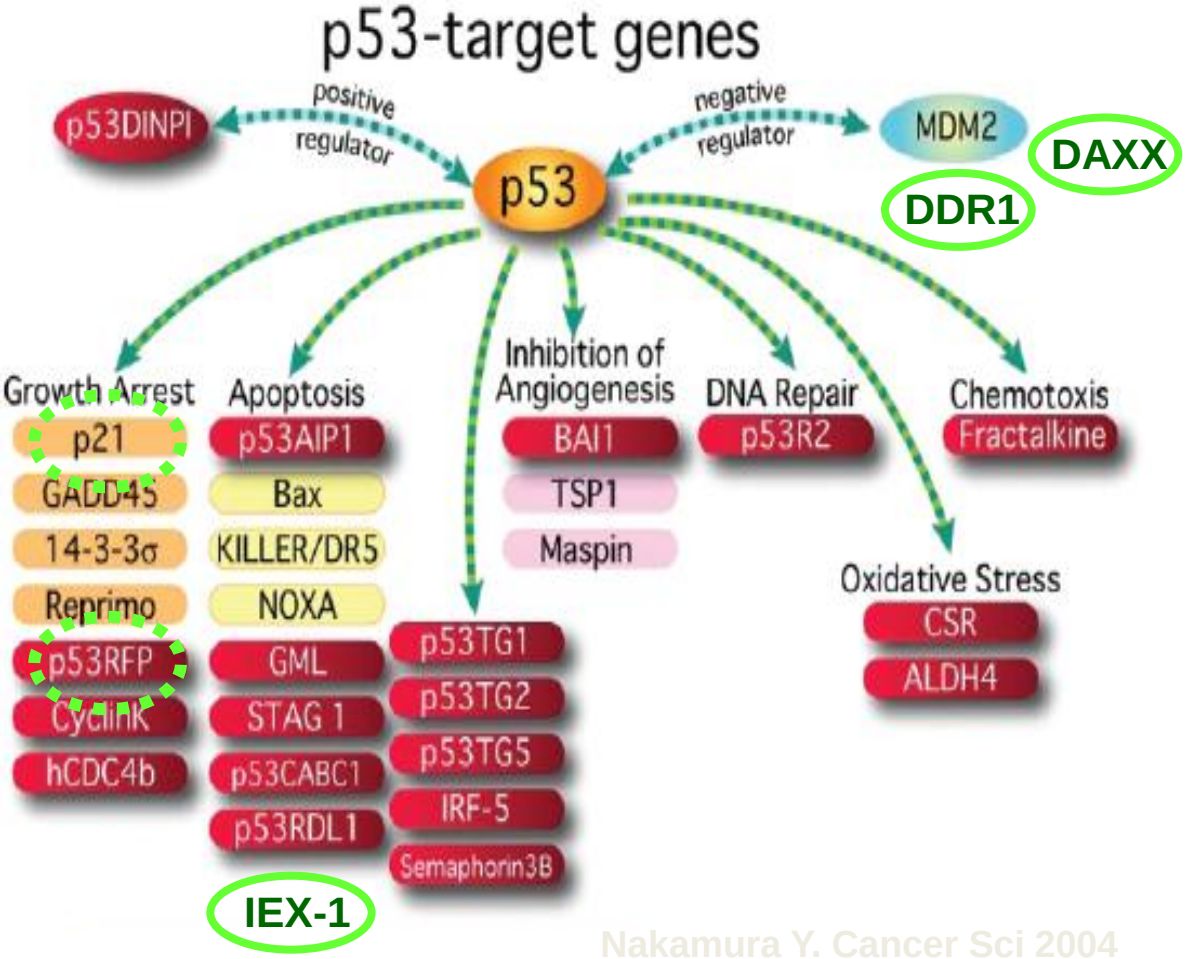
Current evidence suggests that a correlation exists between the capacity to perform excision repair of UV-induced DNA damage and maximum lifespan in different species. Preliminary evidence has also indicated differences of DNA repair capacities in lymphocytes of several strains of mice congenic at the H-2 locus. It is known that the H-2 system influences maximum lifespan potential in mice. In the present studies excision repair of UV-induced DNA damage, but not gamma-induced damage, was found to correlate the mean survival in the adult inbred mouse strains NZB and CBA, using PHA stimulated splenic lymphocytes. Furthermore, in (NZB X CBA)F2 hybrid with adult progeny the level of DNA repair of UV-induced damage corresponded to the H-2 allele (H-2d/2d from NZB or H-2b/2b from CBA) inherited from the parental strain. These studies suggest the possibility of a tricornered relationship between the main histocompatibility complex, one form of DNA repair, and lifespan within the species.

1: Tissue Antigens. 1979 Oct;14(4):336-42.

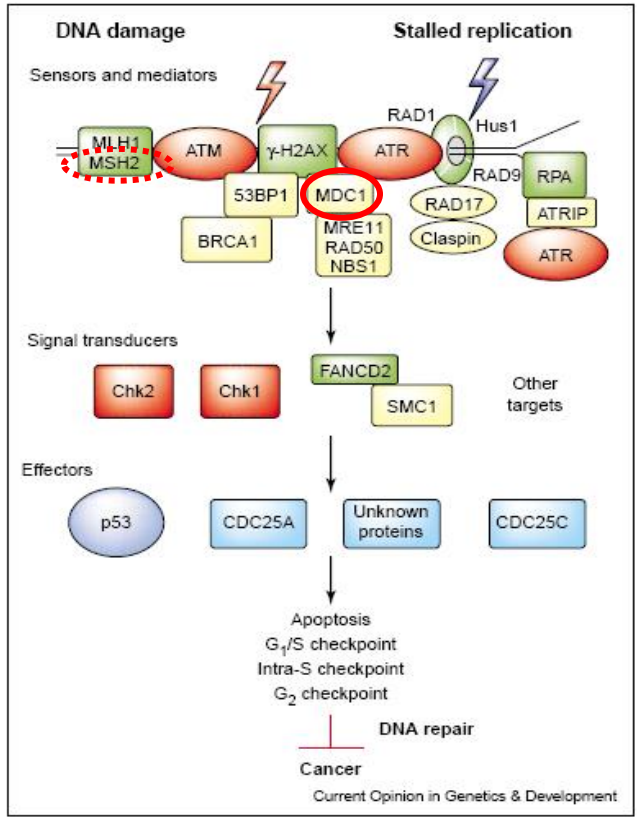
Influence of genes associated with the main histocompatibility complex on deoxyribonucleic acid excision repair capacity and bleomycin sensitivity in mouse lymphocytes.

Walford RL, Bergmann K.

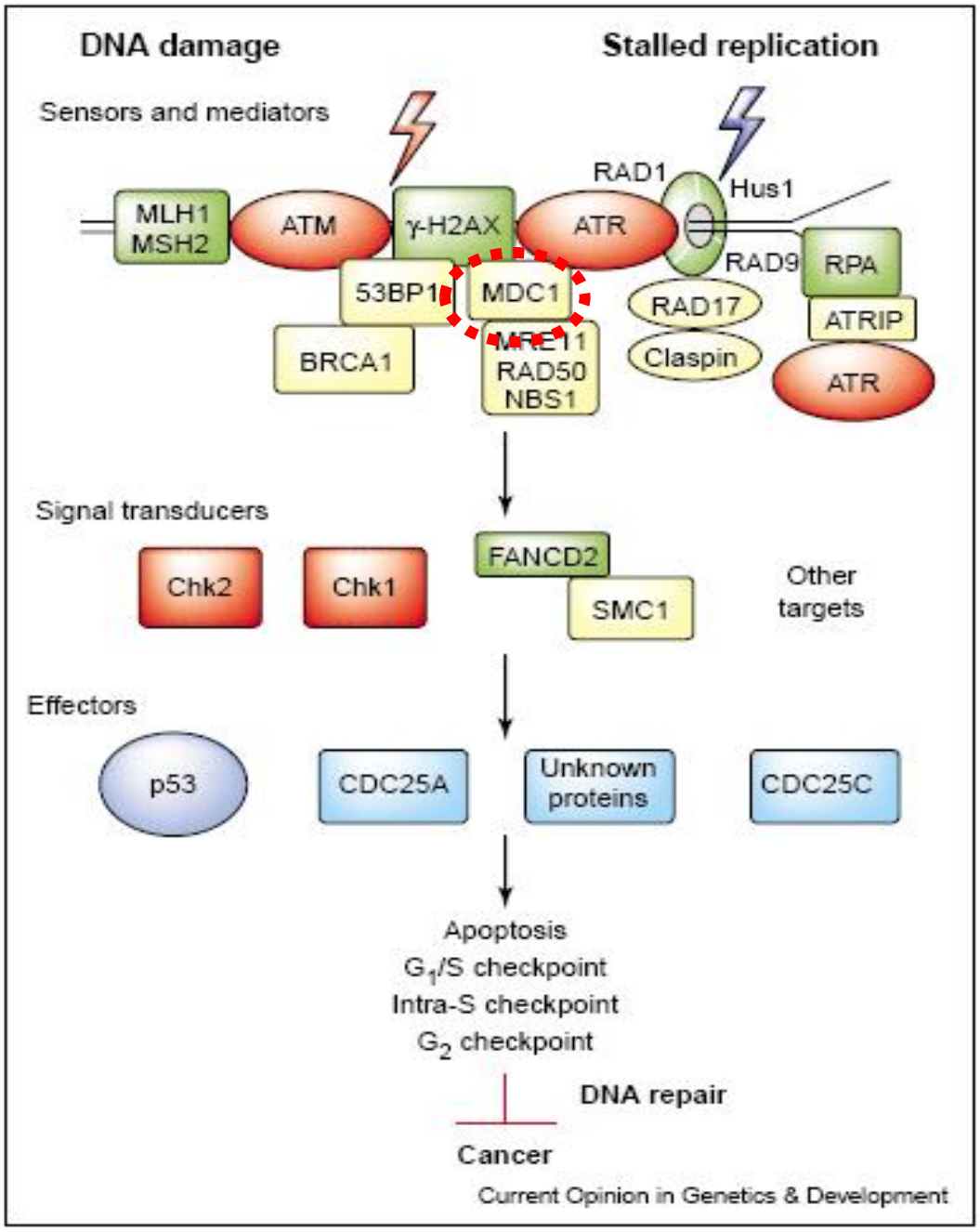
In sets of mice congenic at H-2 and upon two backgrounds, and selected according to known differences in strain-specific lifespans, DNA repair efficiency in spleen cells was compared by two techniques: excision repair capacity following UV-irradiation, and bleomycin sensitivity. Significant differences between certain congenic partner sets were noted with both techniques, suggesting that the main histocompatibility complex influences DNA repair capacity.



Nakamura Y. Cancer Sci 2004



Motoyama, 2004 (www)



DNA damage checkpoint machinery

In response to DNA damage, ATM and ATR phosphorylate histone H2AX and thereby facilitate the recruitment and phosphorylation of mediators such as MDC1, 53BP1, BRCA1, and the MRE11–RAD50–NBS1 complex. Stalling of the DNA replication fork results in the recruitment of the ATR–ATRIP complex by RPA. In turn, the formation of nuclear foci of mediator complexes promotes transmission of the DNA damage signal to downstream targets such as Chk1, Chk2, FANCD2, and SMC1. The PCNA-like RAD1–RAD9–Hus1 complex, the RFC-like RAD17, and Claspin may collaborate in checkpoint regulation by detecting different aspects of a DNA replication fork. The mismatch repair proteins MLH1 and MSH also implicate in the activation of ATM-Chk2 pathway. The kinases Chk1 and Chk2 phosphorylate effectors such as p53, CDC25A, and CDC25C and thereby delay cell cycle progression or induce senescence or apoptosis via activation of the G1–S, intra-S, or G2 cell cycle checkpoints. Thus, these DNA damage checkpoint mechanisms cooperate with DNA repair machinery to suppress genomic instability and cancer.

Motoyama, 2004 (www)

MDC1 is required for the intra-S-phase DNA damage checkpoint

Michal Goldberg*†, Manuel Stucki*†, Jacob Falck‡, Damien D'Amours*, Dinah Rahman§, Darryl Pappin§, Jiri Bartek‡ & Stephen P. Jackson*

** The Wellcome Trust/Cancer Research UK Institute of Cancer and Developmental Biology and Department of Zoology, University of Cambridge, Cambridge CB2 1QR, UK*
‡ Institute of Cancer Biology, Danish Cancer Society, Strandboulevarden 49, DK-2100, Copenhagen, Denmark
§ Department of Proteomics, Imperial College School of Medicine, Hammersmith Campus, London W12 0NN, UK
† These authors contributed equally to this work

MDC1 is coupled to activated CHK2 in mammalian DNA damage response pathways

Zhenkun Lou, Katherine Minter-Dykhouse, Xianglin Wu & Junjie Chen

Department of Oncology, Mayo Foundation, Rochester, Minnesota 55905, USA

Breast Cancer Res Treat (2011) 126:601–607
DOI 10.1007/s10549-010-0960-6

PRECLINICAL STUDY

Mediator of DNA damage checkpoint protein 1 (MDC1) expression as a prognostic marker for nodal recurrence in early-stage breast cancer patients treated with breast-conserving surgery and radiation therapy

Akshar N. Patel · Sharad Goyal · Hao Wu · Devora Schiff · Meena S. Moran · Bruce G. Haffty

Table 3 Univariate analysis of prognostic factors and outcomes (*P* values)

Prognostic factor	Nodal failure	OS
MDC1 expression	0.05	0.45
ER status	0.63	0.37
PR status	0.77	0.92
HER2 status	0.99	0.77
Race	0.55	0.69
Triple negative status	0.24	0.48
Age >50 years	0.98	0.38
T stage	0.47	0.06
N stage	0.99	0.01

OS overall survival, *MDC1* mediator of DNA damage checkpoint protein 1, *ER* estrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor 2, *T* tumor, *N* node

Table 4 Multivariate analysis of prognostic factors and outcomes (*p* values)

Prognostic factor	Nodal failure	OS
MDC1 expression	0.99	0.46
ER status	0.99	0.94
PR status	0.99	0.91
HER2 status	0.99	0.18
Race	1	0.07
Age > 50 years	0.99	0.16
T stage	1	0.10
N stage	0.99	0.84

OS overall survival, *MDC1* mediator of DNA damage checkpoint protein 1, *ER* estrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor 2, *T* tumor, *N* node

nature cell biology

DAXX gene is within the HLA complex.

Critical role for Daxx in regulating Mdm2

Jun Tang[1,7], Li-Ke Qu[1,7], Jianke Zhang[2], Wenge Wang[3], Jennifer S. Michaelson[4], Yan Y. Degenhardt[5,6], Wafik S. El-Deiry[3] and Xiaolu Yang[1,8]

The tumour suppressor p53 induces apoptosis or cell-cycle arrest in response to genotoxic and other stresses[1,2]. In unstressed cells, the anti-proliferative effects of p53 are restrained by mouse double minute 2 (Mdm2), a ubiquitin ligase (E3) that promotes p53 ubiquitination and degradation[3]. Mdm2 also mediates its own degradation through auto-ubiquitination. It is unclear how the cis- and trans-E3 activities of Mdm2, which have opposing effects on cell fate, are differentially regulated. Here, we show that death domain-associated protein (Daxx)[4] is required for Mdm2 stability. Downregulation of Daxx decreases Mdm2 levels, whereas overexpression of Daxx strongly stabilizes Mdm2. Daxx simultaneously binds to Mdm2 and the deubiquitinase Hausp, and it mediates the stabilizing effect of Hausp on Mdm2. In addition, Daxx enhances the intrinsic E3 activity of Mdm2 towards p53. On DNA damage, Daxx dissociates from Mdm2, which correlates with Mdm2 self-degradation. These findings reveal that Daxx modulates the function of Mdm2 at multiple levels and suggest that the disruption of the Mdm2–Daxx interaction may be important for p53 activation in response to DNA damage.

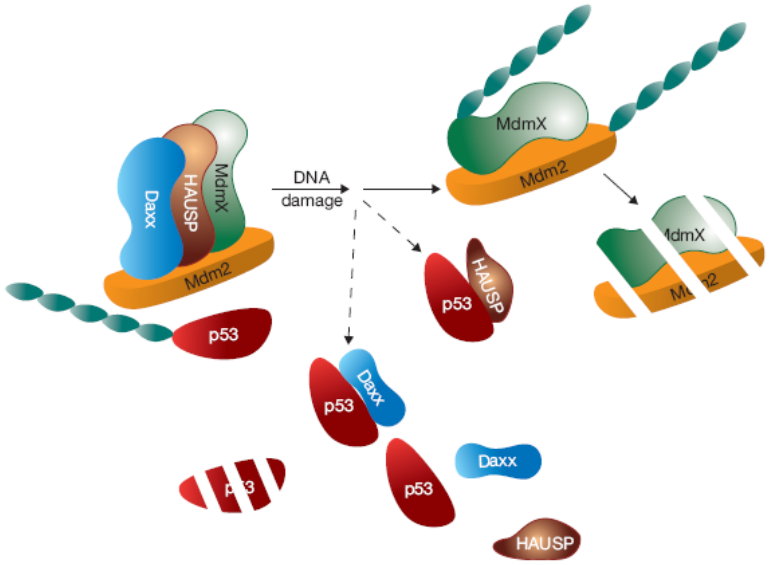


Figure 1 Under non-stress conditions, Daxx associates with HAUSP and Mdm2, which results in stabilization of Mdm2 and MdmX and direction of Mdm2 ligase activity toward p53 that, in turn, leads to p53 ubiquitination and degradation. In response to DNA damage and phosphorylation, dissociation of HAUSP, Daxx and p53 from Mdm2 occurs and the resulting Mdm2–MdmX complex is auto-ubiquitinated and degraded. The remaining components (HAUSP, Daxx and p53) may rearrange to form several hypothetical complexes, leading to different p53 functions.

Oncogene (1998) 16, 2479–2487

http://www.stockton-press.co.uk/onc

The proliferation-associated early response gene p22/PRG1 is a novel p53 target gene

Heiner Schäfer[1,3], Anna Trauzold[1,3], Thorsten Sebens[1], Wolfgang Deppert[2], Ulrich R Fölsch and Wolfgang E Schmidt[1]

[1]Laboratory of Molecular Gastroenterology, 1st Department of Medicine, Christian-Albrechts-University of Kiel, Schittenhelmstr. 12 24105 Kiel, Germany and [2]Heinrich Pette Institute for Experimental Virology & Immunology, University of Hamburg, Germany

IEX1 / IER3 gene is within the HLA complex.

ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biochemical and Biophysical Research Communications 323 (2004) 1293–1298

BBRC

www.elsevier.com/locate/ybbrc

Immediate early gene X-1 interacts with proteins that modulate apoptosis

Rajiv Kumar[a,*], Ward Lutz[a], Elena Frank[a], Hee-Jeong Im[b]

[a] *Division of Nephrology and Hypertension, Departments of Internal Medicine, Biochemistry and Molecular Biology, Mayo Clinic and Foundation, Rochester, MN 55905, USA*
[b] *Department of Biochemistry, Rush University Medical Center, Chicago, IL 60612, USA*

Apoptosis 2003; 8: 11–18

Roles of the stress-induced gene IEX-1 in regulation of cell death and oncogenesis

M. X. Wu

Department of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, Boston, MA 02118, USA

Oncogene (2002) 21, 6819–6828

www.nature.com/onc

Synergistic and opposing regulation of the stress-responsive gene IEX-1 by p53, c-Myc, and multiple NF-κB/rel complexes

Yan-Hong Huang[1], Jim Yujin Wu[2], Yujin Zhang[1] and Mei X Wu*[,1,3]

[1]Department of Pathology, Baylor College of Medicine, Houston, Texas, TX 77030, USA; [2]Shionogi BioResearch, Lexington, Massachusetts, MA 02420, USA

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 277, No. 17, Issue of April 26, pp. 14612–14621, 2002
Printed in U.S.A.

Divergent Regulation of the Growth-promoting Gene *IEX-1* by the p53 Tumor Suppressor and Sp1*

Received for publication, September 28, 2001, and in revised form, February 13, 2002
Published, JBC Papers in Press, February 13, 2002, DOI 10.1074/jbc.M109414200

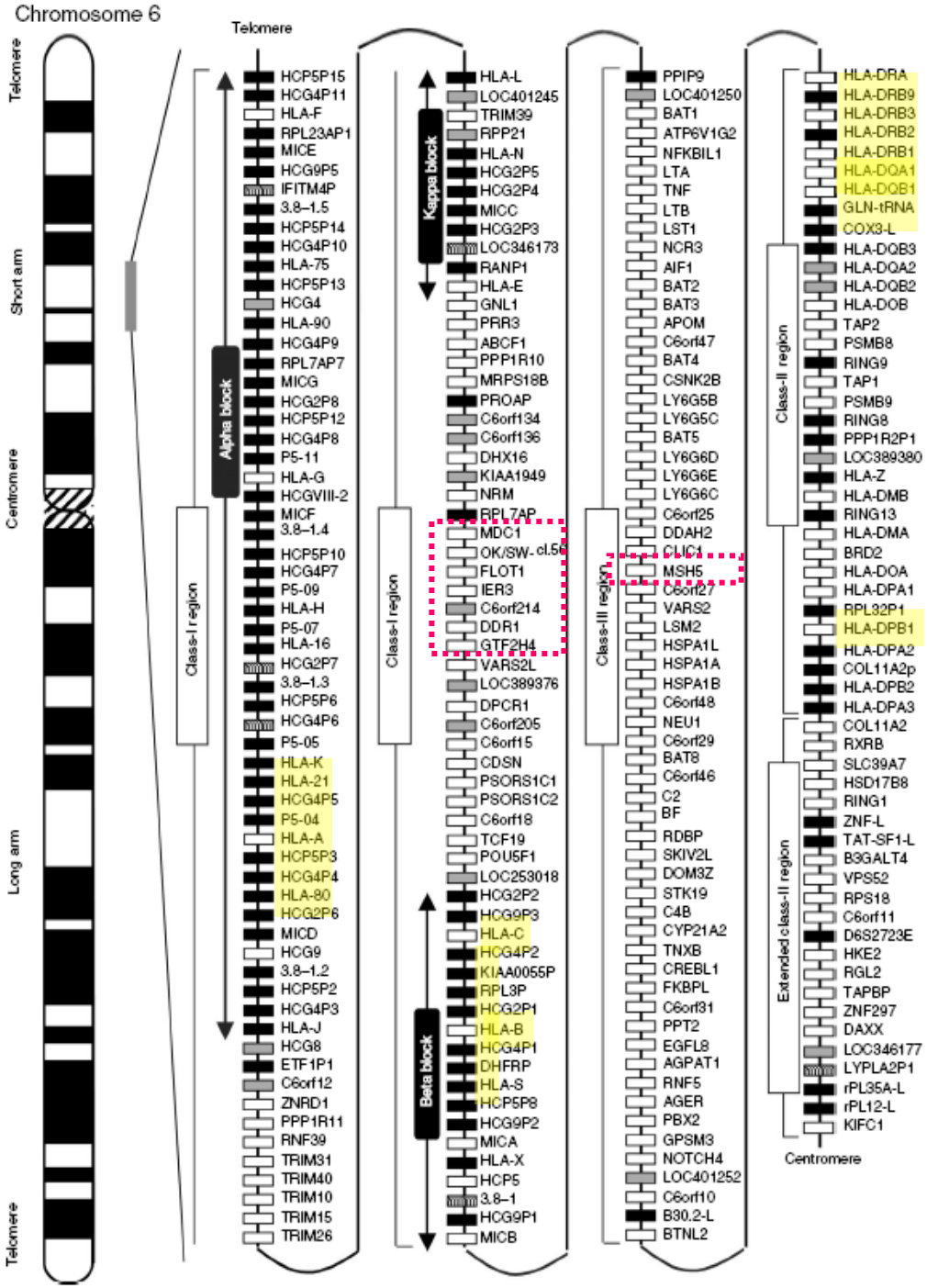
Hee-Jeong Im‡, Mark R. Pittelkow§, and Rajiv Kumar‡¶

From the Departments of ‡Internal Medicine, Biochemistry, and Molecular Biology and §Dermatology, Mayo Clinic and Foundation, Rochester, Minnesota 55905

FIU FLORIDA INTERNATIONAL UNIVERSITY

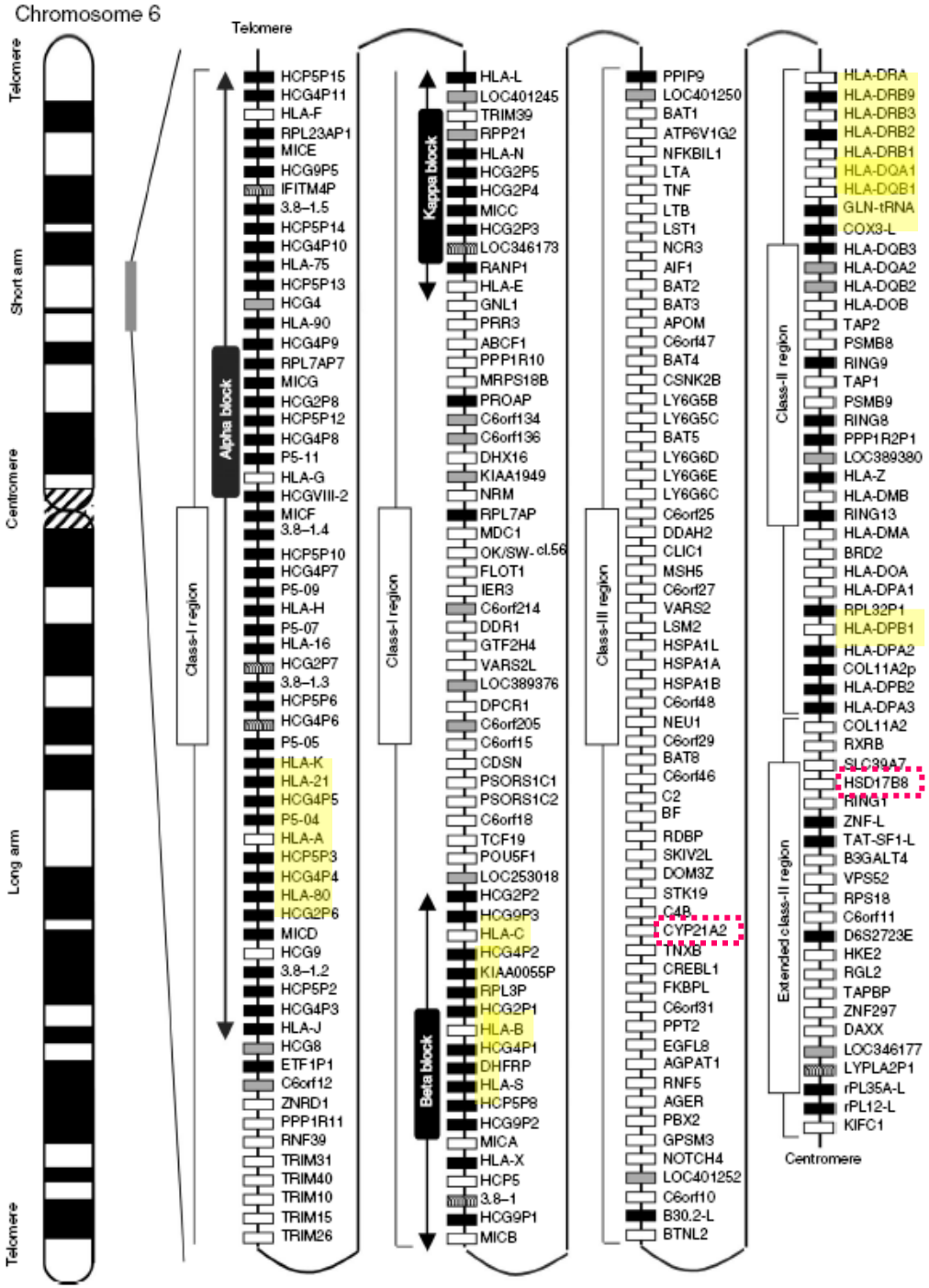
(in the 1980s, we did not know much about non-HLA genes)

But now we do!



Shiina et al, 2004 (www)

CYP21A2 & Adrenal Sex Steroids



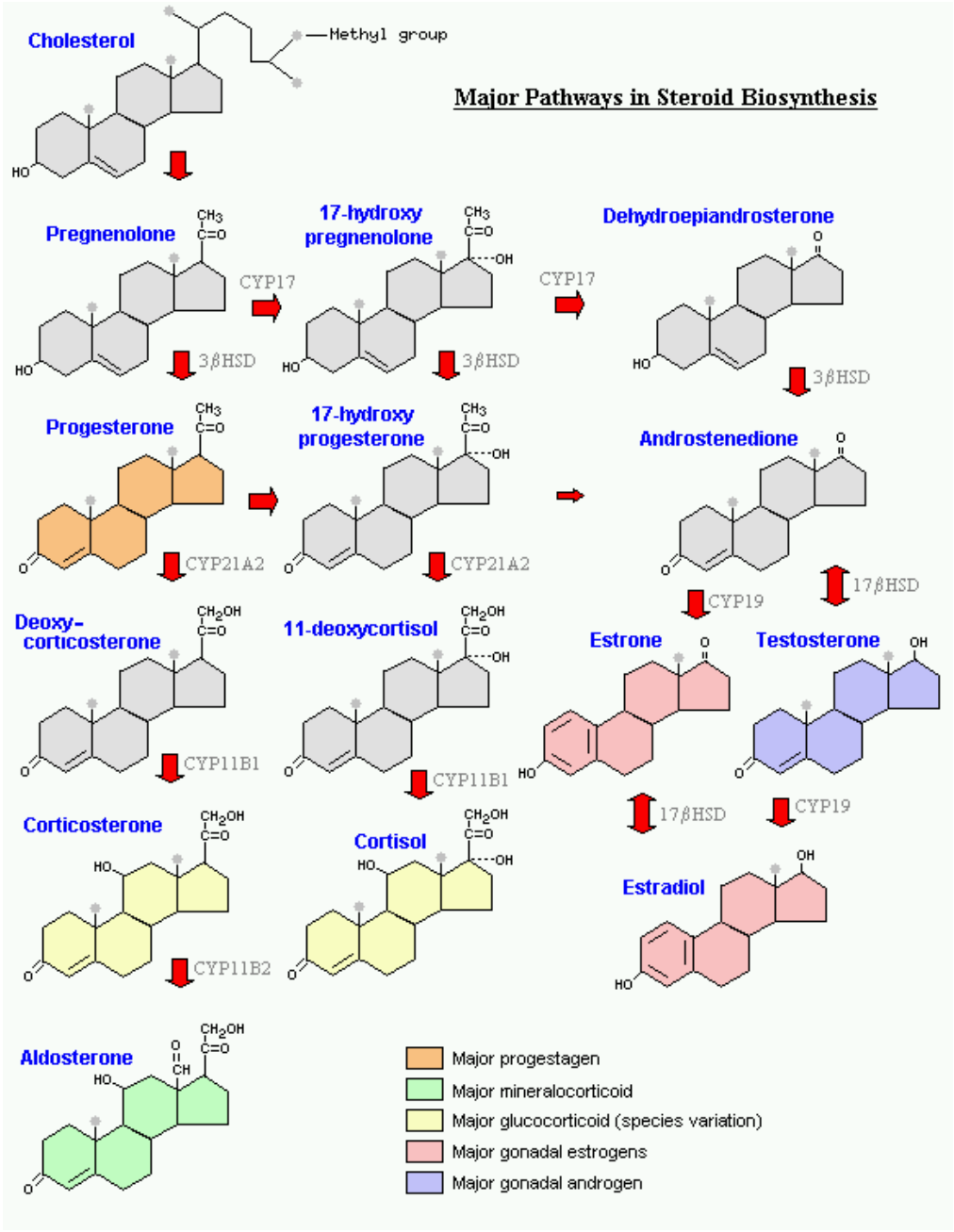
Shiina et al, 2004 (www)

CYP21A2 & Adrenal Sex Steroids

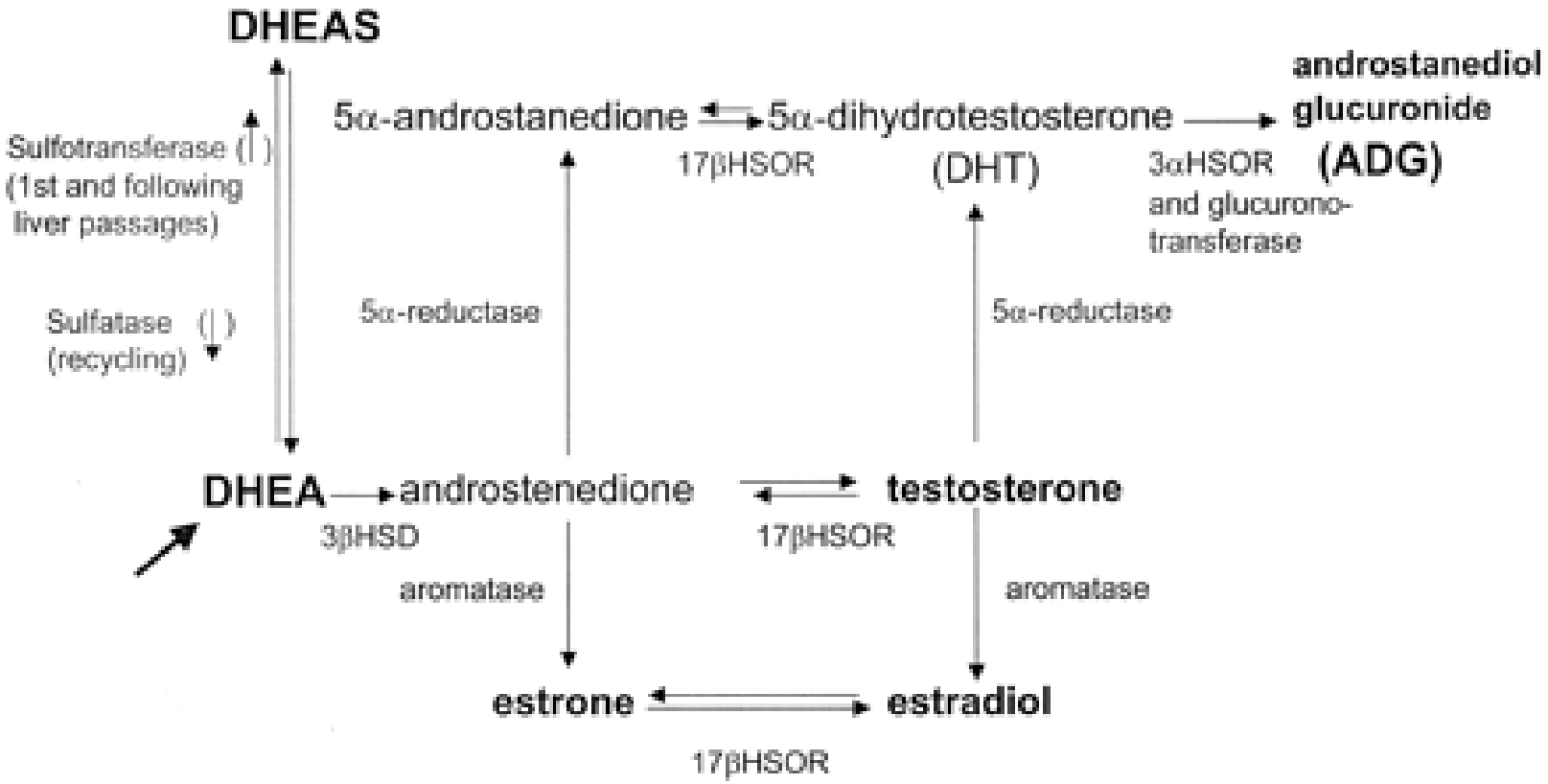
Estrogen exposure increases the risk for breast cancer

After menopause, the only source of endogenous estrogen is adrenal sex steroids

Adrenal sex steroid levels show associations with risk and prognosis


Cholesterol
Methyl group
Major Pathways in Steroid Biosynthesis
Pregnenolone
CYP17
17-hydroxy pregnenolone
CYP17
Dehydroepiandrosterone
3βHSD
3βHSD
3βHSD
Progesterone
17-hydroxy progesterone
Androstenedione
CYP21A2
CYP21A2
17βHSD
CYP19
Deoxy-corticosterone
11-deoxycortisol
Estrone
Testosterone
CYP11B1
CYP11B1
17βHSD
CYP19
Corticosterone
Cortisol
Estradiol
CYP11B2
Aldosterone
Major progestagen
Major mineralocorticoid
Major glucocorticoid (species variation)
Major gonadal estrogens
Major gonadal androgen

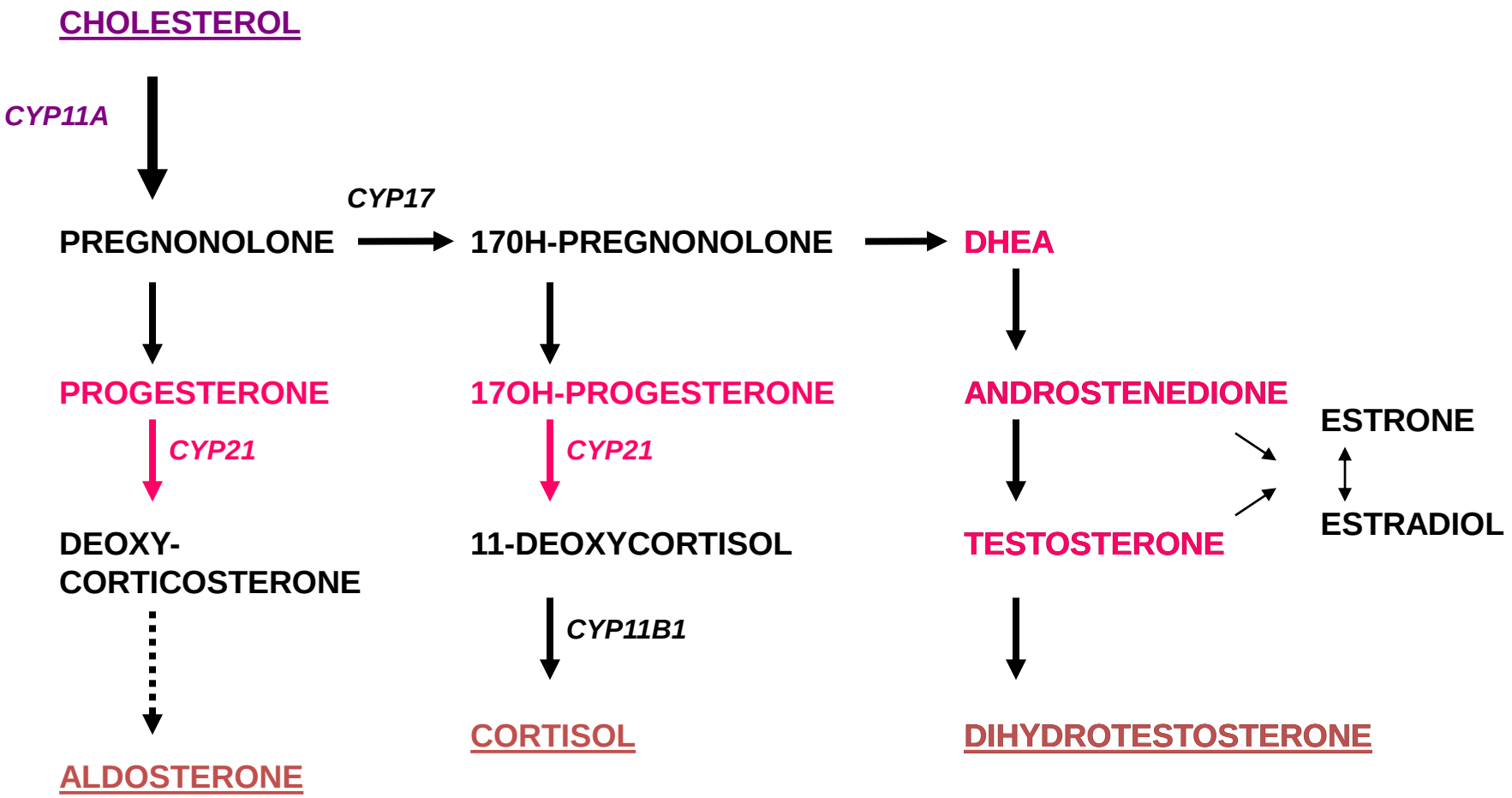
DHEA(S) METABOLISM



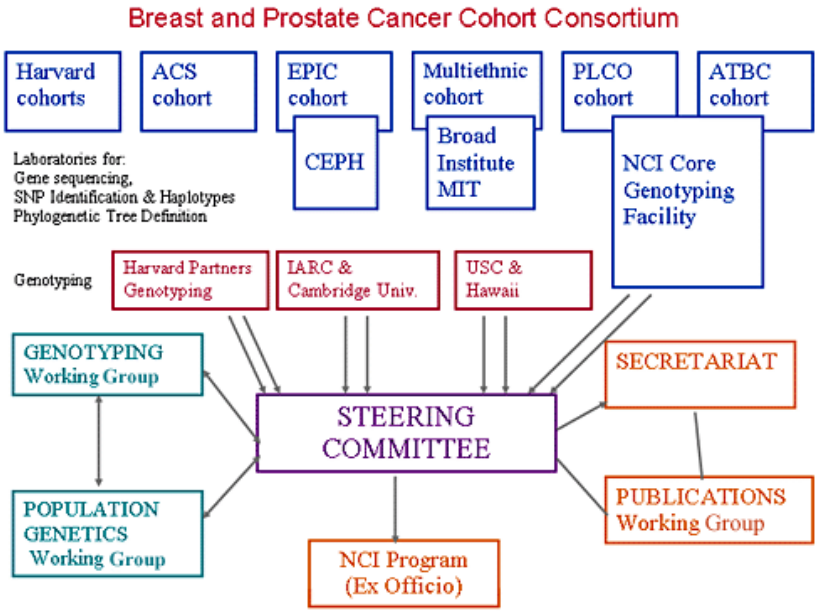
Legraine, JCEM 2000 (www)

http://jcem.endojournals.org/cgi/content/full/85/9/3208/F1

Adrenal Steroid Biosynthesis Pathway



FIU FLORIDA INTERNATIONAL UNIVERSITY



National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov

Epidemiology and Genomics Research
Cancer Control and Population Sciences

About the Program | Key Research | Funding Opps | Our Grantees | Workshop Reports | More Resources

Search: GO

- Cancer Control and Population Sciences Home
- Epidemiology and Genomics Research Home
- Consortia (EGRP Facilitated and Funded)
- Cohort Consortium
- **Breast and Prostate Cancer Cohort (BPC3) Consortium:**
  - BPC3 Home
  - Original Research Plan
  - Current Research Plan
  - Consortium Members
  - Genotyping Resources
  - Publications
  - **Archived BPC3 Information**
  - Related Grants

Breast and Prostate Cancer and Hormone-Related Gene Variant Study

Gene Resequencing

Extensive resequencing is taking place on 75 candidate genes at three Genomics Centers:

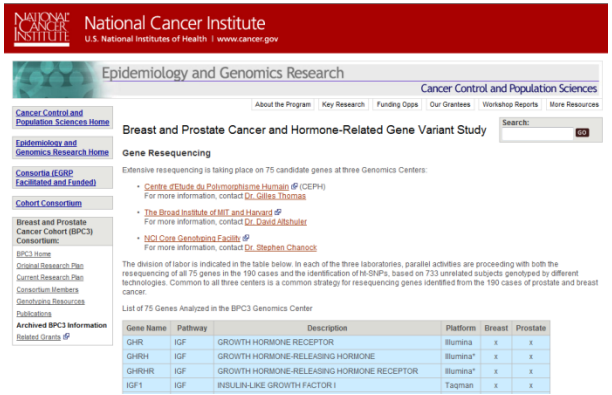
- Centre d'Etude du Polymorphisme Humain (CEPH)
  For more information, contact Dr. Gilles Thomas
- The Broad Institute of MIT and Harvard
  For more information, contact Dr. David Altshuler
- NCI Core Genotyping Facility
  For more information, contact Dr. Stephen Chanock

The division of labor is indicated in the table below. In each of the three laboratories, parallel activities are proceeding with both the resequencing of all 75 genes in the 190 cases and the identification of ht-SNPs, based on 733 unrelated subjects genotyped by different technologies. Common to all three centers is a common strategy for resequencing genes identified from the 190 cases of prostate and breast cancer.

List of 75 Genes Analyzed in the BPC3 Genomics Center

Gene Name	Pathway	Description	Platform	Breast	Prostate
GHR	IGF	GROWTH HORMONE RECEPTOR	Illumina	x	x
GHRH	IGF	GROWTH HORMONE-RELEASING HORMONE	Illumina*	x	x
GHRHR	IGF	GROWTH HORMONE-RELEASING HORMONE RECEPTOR	Illumina*	x	x
IGF1	IGF	INSULIN-LIKE GROWTH FACTOR I	Taqman	x	x

FIU | FLORIDA INTERNATIONAL UNIVERSITY



MPO	replication	MYELOPEROXIDASE	Taqman	x	x
ACTHR	steroid	ADRENOCORTICOTROPIC HORMONE RECEPTOR	Taqman	x	x
ACVR1	steroid	ACTIVIN A RECEPTOR, TYPE I	Illumina	x	x
ACVR2	steroid	ACTIVIN A RECEPTOR, TYPE II	Illumina	x	x
AKR1C3	steroid	ALDO-KETO REDUCTASE FAMILY 1, MEMBER C3	Illumina	x	x
AR	steroid	ANDROGEN RECEPTOR	Taqman	x	x
CGA	steroid	CHORIONIC GONADOTROPIN	Illumina	x	x
COMT	steroid	CATECHOL - O -METHYL TRANSFERASE	Taqman	x	x
CYP11A	steroid		Taqman	x	x
CYP17	steroid		Taqman	x	x
CYP19	steroid		Taqman	x	x
CYP1A1	steroid		Taqman	x	x
CYP1A2	steroid		Taqman	x	x
CYP1B1	steroid		Taqman	x	x
CYP3A4	steroid		Illumina	x	x
ESR1	steroid	ESTROGEN RECEPTOR ALPHA	Illumina	x	x
ESR2	steroid	ESTROGEN RECEPTOR BETA	Taqman*	x	x
FSHB	steroid	FOLLICLE-STIMULATING HORMONE BETA UNIT	Taqman	x	x
FSHR	steroid	FOLLICLE-STIMULATING HORMONE RECEPTOR	Illumina	x	x
HSD17B1	steroid	17-BETA-HYDROXYSTEROID DEHYDROGENASE 1	Taqman	x	x
HSD17B2	steroid	17-BETA-HYDROXYSTEROID DEHYDROGENASE 2	Taqman	x	x
HSD17B3	steroid	17-BETA-HYDROXYSTEROID DEHYDROGENASE 3	Illumina	x	x
HSD17B4	steroid	17-BETA-HYDROXYSTEROID DEHYDROGENASE 4	Illumina	x	x
HSD3B1/2	steroid	3-BETA-HYDROXYSTEROID DEHYDROGENASE	Illumina	x	x
INHA	steroid	INHIBIN, ALPHA	Illumina	x	x
INHBA	steroid	INHIBIN, BETA A	Illumina	x	x
INHBB	steroid	INHIBIN, BETA B	Illumina	x	x
LHB	steroid	LUTEINIZING HORMONE, BETA	Illumina	x	x
LHCGR	steroid	LUTEINIZING HORMONE/CHORIOGONADOTROPIN RECEPTOR	Illumina	x	x
PGR	steroid	PROGESTERONE RECEPTOR	Illumina*	x	x
POMC	steroid	PROOPIOMELANOCORTIN	Illumina	x	x
PRL	steroid	PROLACTIN	Illumina	x	x
PRLR	steroid	PROLACTIN RECEPTOR	Illumina	x	x
SHBG	steroid	SEX HORMONE-BINDING GLOBULIN	Illumina	x	x
SRD5A1	steroid	STEROID 5-ALPHA-REDUCTASE 1	Illumina	x	x

FIU FLORIDA INTERNATIONAL UNIVERSITY

CYP21A2 gene contains common mutations that cause CAH

In heterozygosity, these mutations increase DHEAS levels

This gene is not included in any GWAS chip

This gene has never been examined in breast cancer susceptibility studies

Can We Improve Breast Cancer Predictive Models?

Yes, We Can…

- GWAS are not the ultimate tools as frequently thought
- Most relevant markers are not included
- Data analysis for GWAS wastes a lot of data
- Heterozygote advantage, haplotype analysis and interactions are not considered
- HLA complex data is overlooked

Conclusions

- Current knowledge on genetic susceptibility to breast cancer is insufficient
- GWAS has substantially contributed to the field, but is not the ultimate tool
- Increasing the number of SNPs in covered regions is unlikely to improve the situation
- HLA region is a strong candidate to harbor breast cancer susceptibility markers

Take Home Message

- GWAS do not cover the whole genome
- HLA complex contains biologically highly plausible candidate genes for breast cancer susceptibility
- Filling the gaps of GWAS may be more productive than increasing the number of SNPs included in GWAS chips

Acknowledgments

Cardiff, UK	Professor Alan K Burnett Professor Ken I Mills
Birmingham, AL	Dr Wenshuo Shao Dr Richard A Kaslow
Newcastle, UK	Professor Andy G Hall Professor Louise Parker Dr Caroline Relton Rachel MacKay
Hamilton, NJ	Brittany Morrison Charronne Davis Esma Ucisik-Akkaya Dr Thuy Do Dr Grant Gallagher & Dr Eli Mordechai

FIU | FLORIDA INTERNATIONAL UNIVERSITY

Acknowledgments

My current team:
Amy, Sandeep & Malar (& Meredith)

Florida International University
RSCPHSW
EOH

FIU | FLORIDA INTERNATIONAL UNIVERSITY