Linkage Analysis and Complex Traits

Elaine A. Ostrander, Ph.D.

Chief, Cancer Genetics Branch
Head, Section of Comparative Genetics
National Human Genome Research Institute
National Institutes of Health

Linkage-Based Approaches to Finding Susceptibility Genes

- Linkage Analysis Using High Risk Families
- Analysis of Families with Shared Phenotypic Features
- Linkage Studies of Multi-Cancer Families
- Genetic Analysis of Isolated Populations

Linkage-Based Approaches to Finding Susceptibility Genes

- Linkage Analysis Using High Risk Families
- Analysis of Families with Shared Phenotypic Features
- Linkage Studies of Multi-Cancer Families
- Genetic Analysis of Isolated Populations

Prostate Cancer

Most Common Cancer in the U.S. for Men

- *234,460 new cases to be diagnosed in 2006; about 27,000 deaths
- Median age at diagnosis = 68 yrs

Segregation Analysis Suggests Genetic Factors**

- 9% of prostate cancer in men ≤ 85 years
- 43% of prostate cancer in men < 55 years
- Population prevalence 0.3-1.0%, 88% penetrance by age 85

Epidemiology Studies

- Relatives diagnosed ≤ age 65 or ≥ 3 affected first degree relatives = RR of 10.9

Ries et al., 2005 ; Jemal et al., 2006 Carter et al. 1992; Gronberg et al. 1997; Schaid et al. 1998; Cui et al. 2001*

Estimates of Linkage

- Genome-wide scan
  - Testing for linkage between markers and disease state
- LOD score - Log of Odds
  - Do number of recombinants between marker and putative disease locus differ significantly over chance?
  - Underlying model of inheritance
  - LOD score ≥ 3.3 significant
  - Indicate greater then 1000:1 odds in favor of linkage
- NPL - Nonparametric Linkage Analysis
  - Significant allele sharing among affected individuals?
  - No model of inheritance
  - Assessed as *P* value

255 *PROGRESS* Hereditary Prostate Cancer (HPC) Families

- 1,998 blood samples collected
  - 847 affected men, 613 unaffected men, 538 women
- Average of:
  - 7.8 sampled relatives per family
  - 3.3 sampled affected men per family
- Mean age of diagnosis 65.6
- Genome-wide scan
  - 441 microsatellite markers
  - 8.1 cM average spacing

Janer et al., (2003) Prostate 57:309-319

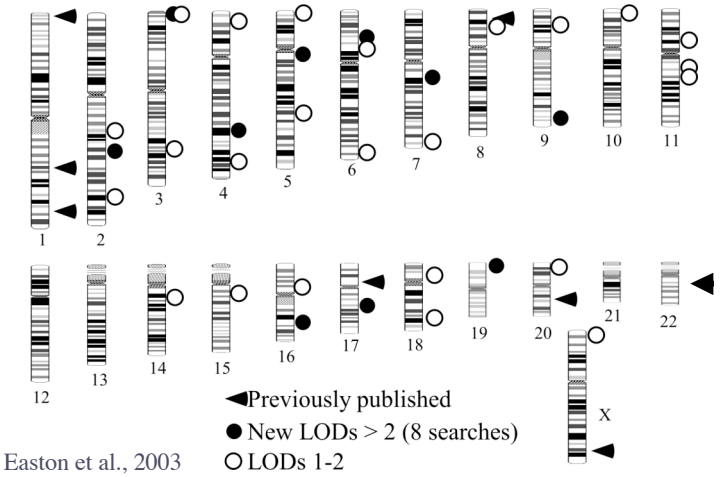
Summary of Linkage Results in 254 *PROGRESS* Families (LOD≥1.9)

Strata (# of families)	Marker	Model	LOD	HLOD
All families (254)	D6S1281	Dominant affected only	**2.36**	**2.51**
		Dominant	**1.70**	**1.93**
	D7S2212	Recessive	**1.55**	**2.25**
Median age of PC onset 56-72 years (214)	D6S1281	Dominant affected only	**3.42**	**3.43**
		Dominant	**2.52**	**2.62**
	D7S2212	Recessive	**1.68**	**2.41**
≥5 sampled affected (26)	D2S1391	Dominant	**2.63**	**2.63**
	D8S1119	Recessive	**2.01**	**2.01**
	D10S1432	Dominant	**1.93**	**2.06**
	D13S285	Recessive	**2.21**	**2.21**

Over 800,000 genotypes completed

Janer et al., (2003) Prostate 57:309-319

Summary of Approximately 15 Individual Prostate Cancer Genome Wide Scans



Easton et al., 2003

Results observed on almost every chromosome.
No chromosomal region with Lod ≥ 2.0 observed by more than one study!

Why So Hard?

- Mapping prostate cancer genes difficult.
  - Late age onset disease
  - Locus heterogeneity
  - High phenocopy rate
  - Variable penetrance
- Each individual research group suffers from a lack of power
  - Finding linkage
  - To reproduce reports

Extreme Locus Heterogeneity in HPC

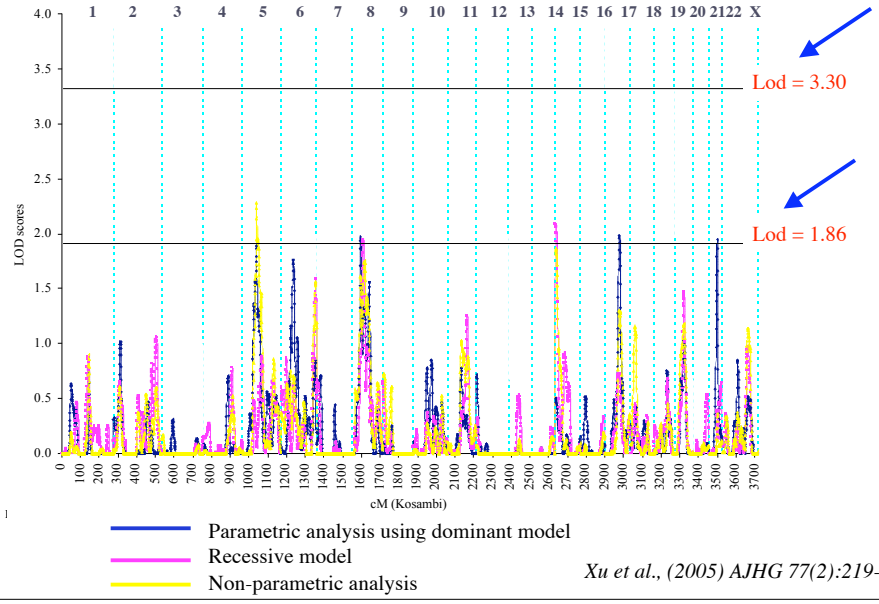
Approaches to overcoming heterogeneity in HPC

- International Consortium of Prostate Cancer Genetics (ICPCG) combined analysis of 1,233 families (Chromosome 22)
- Analysis of families according to clinical features of disease (Chromosome 22)
- Presence of other cancers in HPC families (Chromosome 11)
- Isolated populations with a limited number of founders (Chromosome 7)

ICPCG Resources

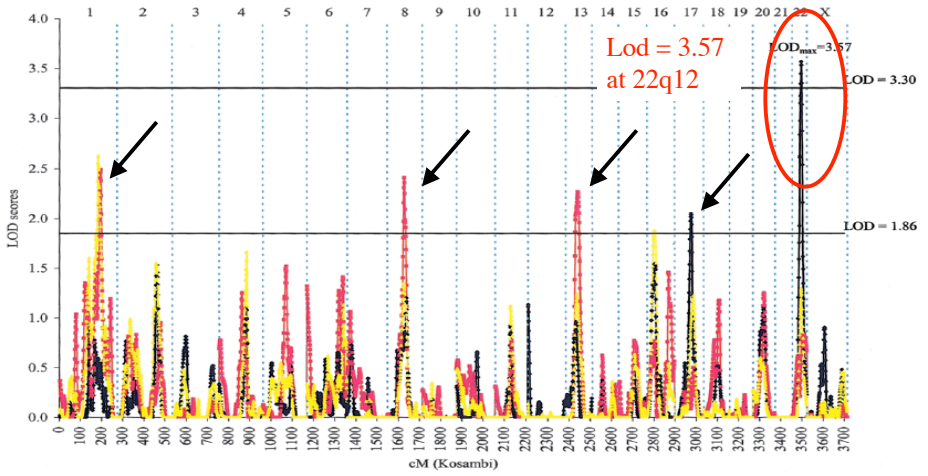
- 2500 multiplex prostate cancer families
  - One of largest family resources in the world for addressing genetic mechanisms cancer susceptibility
  - Over 12,000 DNA samples
  - 6400 sampled affected men
- 11 Research Groups - several institutions
- Data Coordinating Center (DCC)-Wake Forest University
  - Deposition, organization, analysis and dissemination of combined analyses

Combined Genome-Wide Screen Among 1233 ICPCG Families



Xu et al., (2005) AJHG 77(2):219-29

Combined Genome-Wide Screen Among 269 Families with ≥ 5 Affecteds

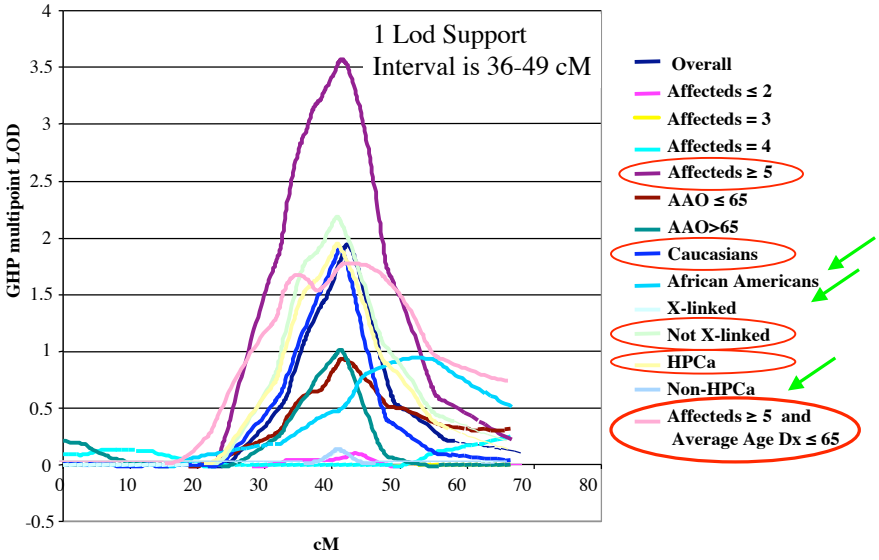


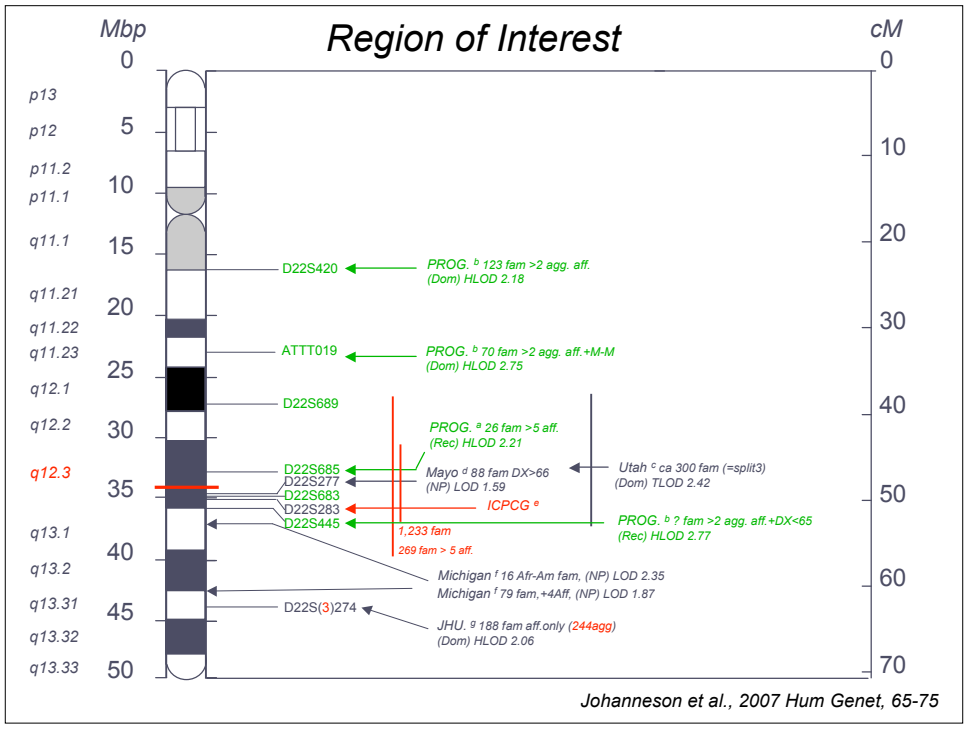
Parametric analysis using a dominant model
Recessive model
Non-parametric analysis

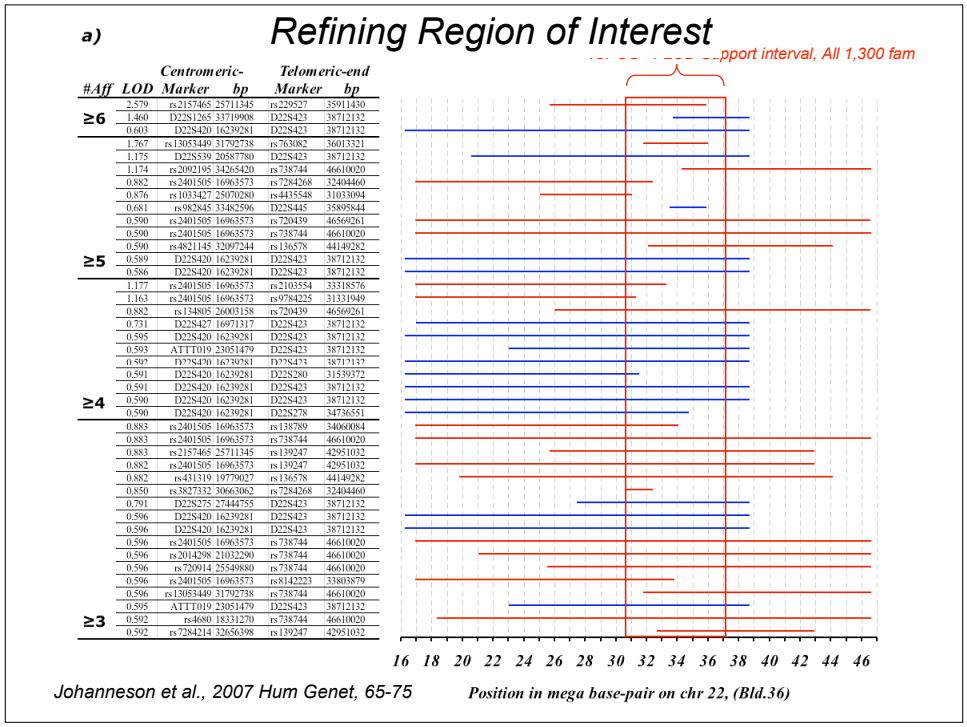
Xu et al., (2005) AJHG 77(2):219-29

Five regions of suggestive linkage (5q12, 8p21, 15q11, 17q21, 22q12) and significant linkage (22q12)

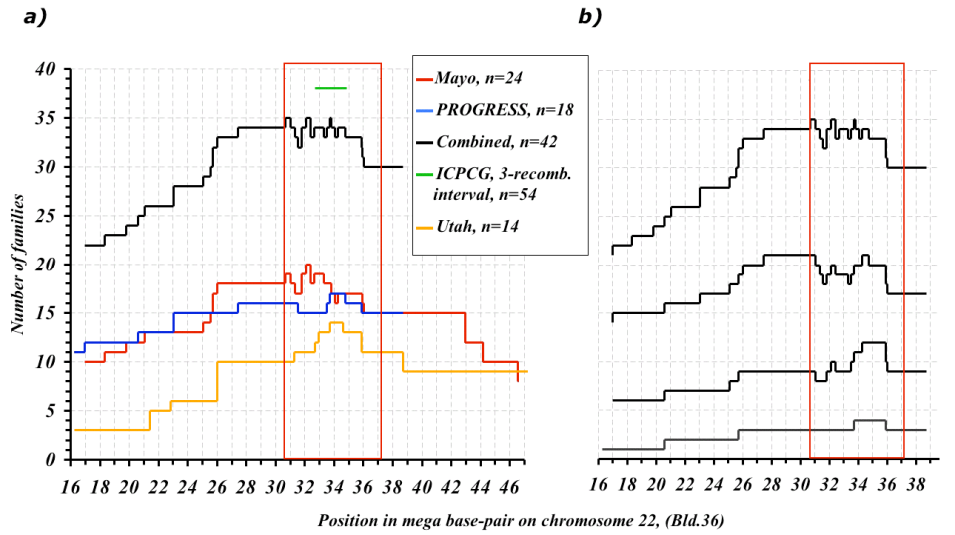
ICPCG Genome wide Scan Chromosome 22 Linkage Results 1139 Families




Region of Interest
Mbp
cM
p13
p12
p11.2
p11.1
q11.1
q11.21
q11.22
q11.23
q12.1
q12.2
q12.3
q13.1
q13.2
q13.31
q13.32
q13.33
D22S420
PROG. b 123 fam >2 agg. aff. (Dom) HLOD 2.18
ATTT019
PROG. b 70 fam >2 agg. aff.+M-M (Dom) HLOD 2.75
D22S689
PROG. a 26 fam >5 aff. (Rec) HLOD 2.21
D22S685
D22S277
D22S683
D22S283
D22S445
Mayo d 88 fam DX>66 (NP) LOD 1.59
ICPCG e
Utah c ca 300 fam (=split3) (Dom) TLOD 2.42
PROG. b ? fam >2 agg. aff.+DX<65 (Rec) HLOD 2.77
1,233 fam
269 fam > 5 aff.
Michigan f 16 Afr-Am fam, (NP) LOD 2.35
Michigan f 79 fam,+4Aff, (NP) LOD 1.87
D22S(3)274
JHU. g 188 fam aff.only (244agg) (Dom) HLOD 2.06
Johanneson et al., 2007 Hum Genet, 65-75

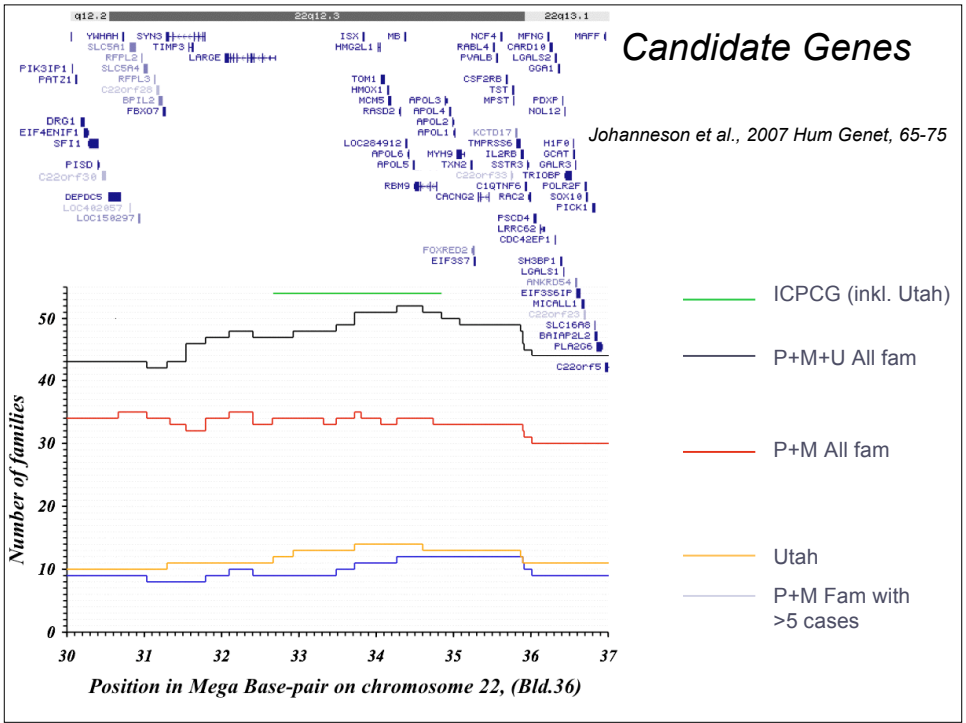

a)
Refining Region of Interest
pport interval, All 1,300 fam
#Aff
LOD
Centromeric-Marker
bp
Telomeric-end Marker
bp
≥6
≥5
≥4
≥3
16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46
Position in mega base-pair on chr 22, (Bld.36)
Johanneson et al., 2007 Hum Genet, 65-75

Peak Region of Linkage



Johanneson et al., 2007 Hum Genet, 65-75

Candidate Genes



Johanneson et al., 2007 Hum Genet, 65-75

Extreme Locus Heterogeneity in HPC

Approaches to overcome the heterogeneity in HPC

- ICPCG combined analysis of 1,233 families
- Analysis of families according to clinical features of disease
- Presence of other cancers in the HPC families
- Isolated populations with a limited number of founders

Mapping Prostate Cancer Aggressiveness Loci

Family Ascertainment

"aggressive families" with ≥3 men with aggressive disease (≥2 genotyped)

PROGESS--123 families met criteria

Definition of Aggressive PC

At least one of the following clinical characteristics:

1) Regional or distant stage pathology, or clinical stage, T3, T4, N1, M1
2) Gleason grade ≥ 7 or poorly differentiated grade
3) Prostate specific antigen at diagnosis ≥ 20 ng/ml
4) Death from metastatic prostate cancer <65 years

PROGRESS Linkage Study for Aggressive Disease

TABLE IV. Summary of Linkage Results Having LOD Scores >2.0 in Subsets of 123 Families With Two or More Men With an Aggressive Prostate Cancer Phenotype

Chromosome	Subset	Position of max, cM	Dom-HLOD	Rec-HLOD	KC-LOD[b]	Flanking markers (cM)	
						Marker (cM)	Marker (cM)
2	No. aff. ≥5	167.9	0.41	1.87	2.10	D2S1353 (162.4)	D2S1776 (170.9)
5	HPC = No	69.2	1.51	1.47	2.06	D5S2500 (68.2)	GATA138B05 (75.9)
6	Dx age ≤58	124.8	1.75	2.16	1.42	D6S474 (117.6)	D6S1040 (127.7)
	HPC = no	61.4	1.18	2.04	1.20	D6S1019 (53.4)	D6S1017 (62.8)
7	No. aff. ≥5	7.4	3.16	0.97	1.80	D7S3056 (7.4)	D7S513 (17.6)
12	Dx age <65	46.2	0.63	1.47	2.25	D12S373 (35.7)	D12S1042 (48.0)
13	No. aff. ≥5	103.6	2.07	0.65	0.96	D13S895 (97.9)	D13S285 (109.5)
20	M to M = no[a]	26.5	2.61	0.66	1.30	ATTC013 (26.4)	D20S604 (32.7)
22	Dx age <65	41.9	0.78	2.77	2.06 (45.8)	D22S683 (35.7)	D22S445 (45.2)
	Dx age (59–70)	15.8	2.32	1.02	1.33	ATTT019 (15.6)	D22S689 (28.1)
	M to M = yes	15.8	2.75	1.79	2.02 (11.1)	ATTT019 (15.6)	D22S689 (28.0)

[a]Suggestive of X-linkage.
[b]Positions (cM) in parentheses refer to the position of the maximum LOD score for a specific model when its position differs from the global maximum LOD score over all three analyses.

Stanford et al., 2006 Prostate, 15:317-25

Extreme Locus Heterogeneity in HPC

Approaches to overcome the heterogeneity in HPC

- ICPCG combined analysis of 1,233 families
- Analysis of families according to disease aggressiveness
- Presence of other cancers in the HPC families
- Isolated populations with a limited number of founders

Prostate Kidney Cancer (KC) Families

- 19 families identified --15 used in this study
- 10 families where KC case = PC case
- 5 families where KC case = 1st degree relative to PC case

- Excluded:
  - Families where KC = 2nd degree relative to PC cases
  - KC patient is not related to any PC cases
  - Wilms tumor family

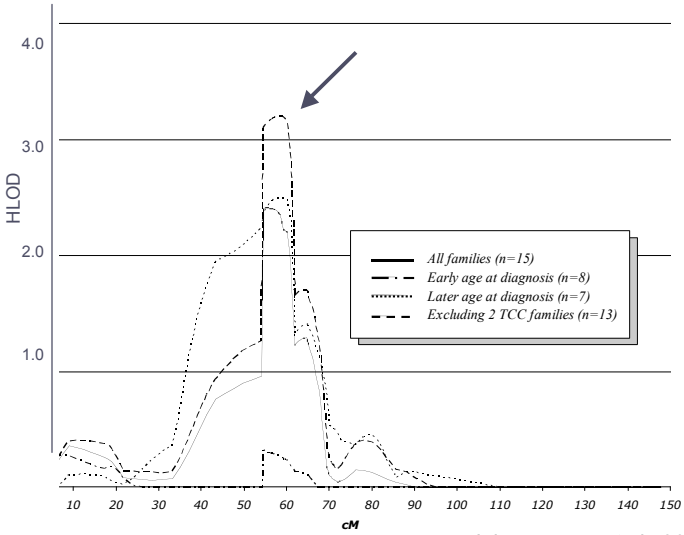
Johannesson et al., 2006, Prostate 15: 732-43

Summary of Linkage Results on Prostate-Kidney Families

Location	cM*	Marker	K&C p-value**	HLOD†	α‡	
1p36.21	29.93	D1S1597	0.02	-	-	
4q21.23	93.48	D4S2361	-	2.099	0.97	11D
7p21.3	17.74	D7S513	0.04	1.905	0.39	AfD
7p14.3	51.79	D7S817	0.03	-	-	
7q34	149.9	D7S1824	0.02	-	-	
8q11.23	67.27	D8S1110	0.04	-	-	
10q26.2	156.27	D10S1223	0.02	-	-	
11q12.1	58.4	D11S1985	0.006	2.591	0.98	11D
12q15	78.06	D12S1294	-	1.742	1.00	
12q23.1	104.13	D12S1300	-	1.920	0.80	11D
15q26.1	90.02	D15S652	-	1.593	1.00	11D
16p12.3	29.97	D16S764	0.02	-	-	
18q22.3	106.81	D18S541	0.02	-	-	

Johannesson et al., 2006, Prostate 15: 732-43

Parametric Multipoint Analysis of Chromosome 11

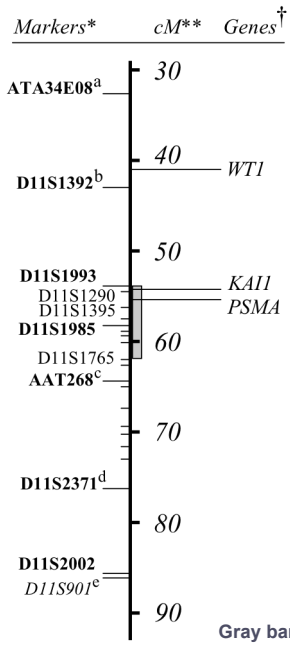


Johannesson et al., 2006, Prostate 15: 732-43

Fine Mapping of 11p11-11q13 Region in HPC-Kidney Families

band	Marker	Mbp*	cM**	HLOD†	α††	K&C *p*-value‡
11p13	**D11S1392**#	34.60	43.16	0.93	0.76	0.04
11p11.2	**D11S1993**	43.57	54.09	1.26	0.72	0.03
11p11.2	D11S1290	44.98	54.50§	3.10	1.00	**0.004**
11p11.12	D11S1395	51.23	56.33§	3.17	1.00	0.005
Centromere	D11S1313	55.99	57.74§	**3.20**	1.00	0.006
11q12.1	D11S4202	58.11	58.36§	3.19	1.00	0.006
11q12.1	**D11S1985**	58.25	58.40	3.19	1.00	0.006
11q12.1	D11S4075	59.26	59.09§	3.19	1.00	0.006
11q12.1	D11S1335	59.29	59.11§	3.19	1.00	0.006
11q12.1	D11S2006	59.47	59.24	3.19	1.00	0.007
11q12.2	D11S4191	59.76	60.09	3.14	1.00	0.008
11q12.2	D11S1765	60.53	61.78	1.64	0.74	0.01
11q12.3	D11S4076	61.11	62.62	1.68	0.74	0.01
11q13.1	**AAT268**	62.82	64.60§	1.70	0.73	0.02
11q13.2	D11S1883	63.12	64.97	1.63	0.73	0.02
11q13.2	D11S913	65.68	67.40	1.24	0.73	0.06
11q13.2	D11S1889	67.06	69.28	0.36	0.43	0.14
11q13.3	D11S987	67.65	69.94	0.23	0.32	0.14
11q13.4	D11S4136	69.31	71.52	0.16	0.26	0.20
11q13.4	D11S4162	70.64	72.75	0.19	0.30	0.20
	D11S2371	73.18	76.13	0.39	0.40	0.20

Johannesson et al., 2006, Prostate 15: 732-43



Candidate Region on Chromosome 11 for Kidney-Prostate Cancer

Gray bar indicates 1 LOD support interval (8cM)

Johannesson et al., 2006, Prostate 15: 732-43

Extreme Locus Heterogeneity in HPC

Approaches to overcome heterogeneity in HPC

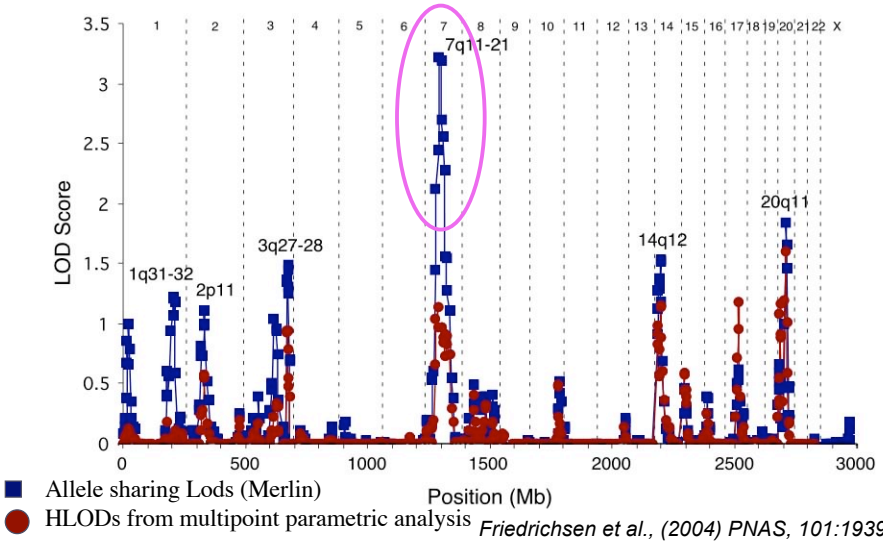
- ICPCG combined analysis of 1,233 families
- Analysis of families according to disease aggressiveness
- Presence of other cancers in the HPC families
- Isolated populations with a limited number of founders

Locus Heterogeneity in HPC

Evaluate families from an isolated population with a limited number of founders

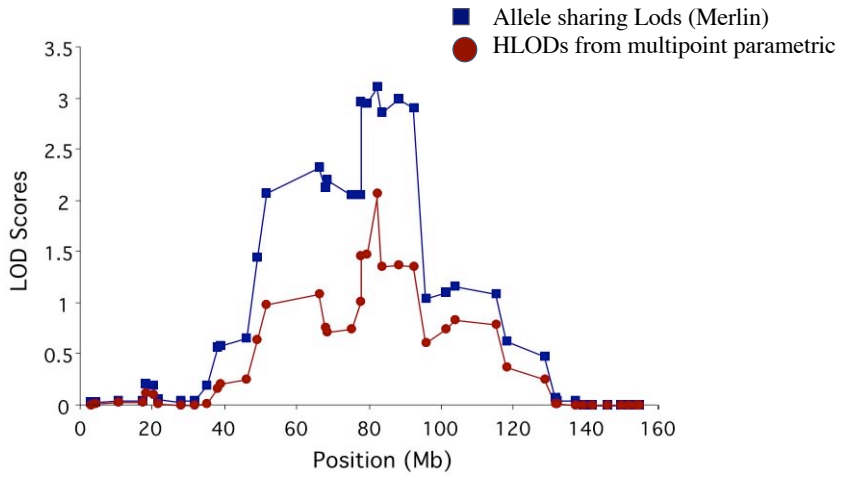
- Americans of (Ashkenazi) Jewish descent
- Predict that only one or two HPC susceptibility genes segregating

Results of Genome-Wide Scan in the 36 Jewish Families Suggest a HPC loci at 7q11-21



Friedrichsen et al., (2004) PNAS, 101:1939-44

Fine Mapping Multipoint Linkage Results Support the 7q11-21 HPC locus



Both Younger and Older Age at Diagnosis Families Contribute to the Result at 7q11-21

	Mean Age at Dx	No. Families	Nonparametric Analysis NPL	*P*	Median No. Affected Men	Median No. Genotyped Affected Men
Younger	< 65	18	2.30	0.011	4.0	2.0
Older	≥ 65	18	3.27	0.0005	4.0	3.0
Total	64.8	36	3.35	0.0004	4.0	3.0

How Much do Jewish Families Account for Original PROGRESS Result?

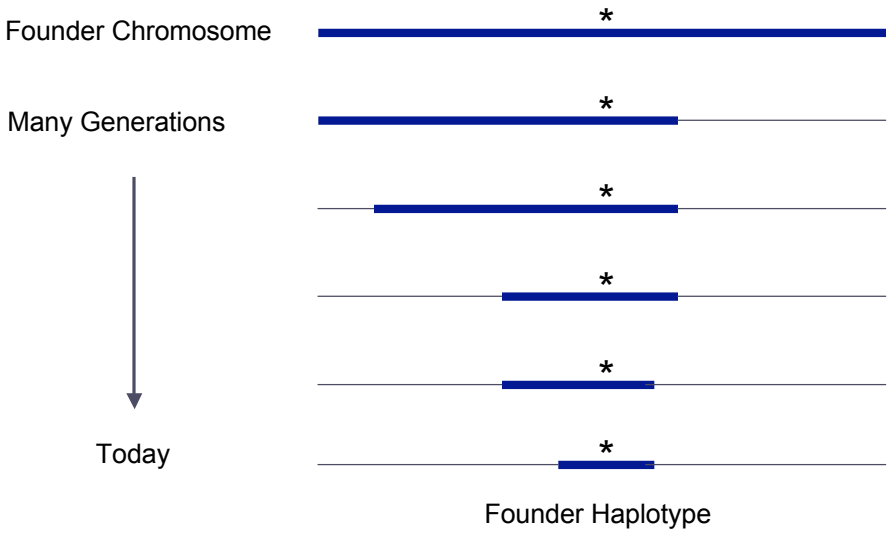
•254 PROGRESS families demonstrate HLOD of 2.25 and NPL of 1.70 (P= 0.038)
•Analysis of 237 non-Jewish Families yield an NPL of 1.11 (P = 0.134)

Majority of PROGRESS results contributed by Jewish families

Strategy for Isolating the Susceptibility Gene

- Identify the founder haplotype surrounding the mutation
  - Founder haplotypes 500 kb – 1 Mb
- Sequence coding regions of genes in regions of shared haplotype
- Initial Approach
  - Focus on minimal recombination regions defined by families
  - Sequence exons of encoded genes
  - Informative SNP every 200 kb on average

What is a Founder Haplotype?



Conclusions

- Prostate cancer genetically heterogenous disease
- Poor replication of linkage results and candidate genes across seemingly similar data sets
- Meta analysis (ICPCG) useful for identifying loci in large families and families with aggressive disease
  - Loci on chromosomes 22 and 11 appear important
  - Multiple other suggestive loci
- Individual dataset analyses supports ICPCG results
- Locus on chromosome 11 important in susceptibility to prostate/kidney cancer, excluding TCC families
- Locus on chromosome 7 important in susceptibility to prostate cancer among Ashkenazi Jewish families

Acknowledgements

PROGRESS Studies

Ostrander Lab- NHGRI-Danielle Friedrichsen, Bo Johannesson, Erika Kwon, Eric Karlins; Seattle-Hawkins DeFrance, Mark Gibbs, Mette Peters, Mariela Langlois

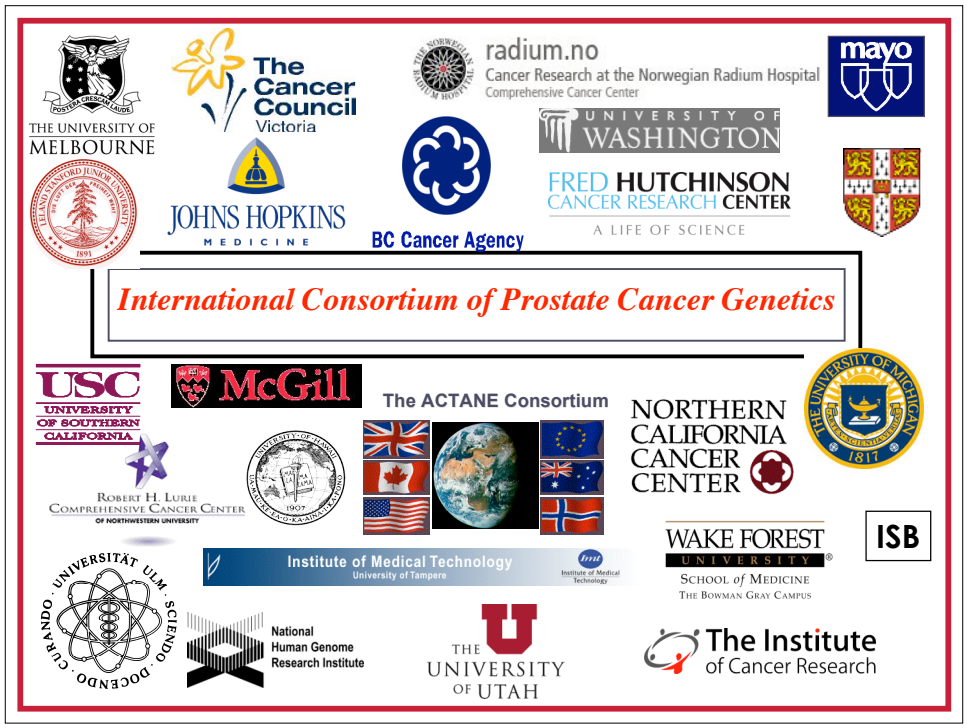
Public Health Sciences-Janet Stanford, Suzanne Kolb, Lori Iwasaki

Institute for Systems Biology -Lee Hood, Marta Janer, Kerry Deutsch

Aggressiveness Studies

Mayo Clinic-Daniel J. Schaid, Shannon K. McDonnell, Erin E. Carlson

Jewish Studies-Wake Forest -Jianfeng Xu, S. Lily Zheng, Bao-li Chang, *Johns Hopkins*- Bill Isaacs, Sarah Isaacs, Katherine Wiley, Pat Walsh


THE UNIVERSITY OF MELBOURNE
The Cancer Council Victoria
radium.no
Cancer Research at the Norwegian Radium Hospital
Comprehensive Cancer Center
mayo
UNIVERSITY OF WASHINGTON
JOHNS HOPKINS MEDICINE
BC Cancer Agency
FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE
International Consortium of Prostate Cancer Genetics
USC UNIVERSITY OF SOUTHERN CALIFORNIA
McGill
The ACTANE Consortium
NORTHERN CALIFORNIA CANCER CENTER
ROBERT H. LURIE COMPREHENSIVE CANCER CENTER
OF NORTHWESTERN UNIVERSITY
ISB
Institute of Medical Technology
University of Tampere
WAKE FOREST UNIVERSITY
SCHOOL of MEDICINE
THE BOWMAN GRAY CAMPUS
UNIVERSITÄT ULM · SCIENDO · DOCENDO · CURANDO
National Human Genome Research Institute
THE UNIVERSITY OF UTAH
The Institute of Cancer Research