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 than 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

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

<table>
<thead>
<tr>
<th>Strata (# of families)</th>
<th>Marker</th>
<th>Model</th>
<th>LOD</th>
<th>HLOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All families (254)</td>
<td>D6S1281</td>
<td>Dominant affected only</td>
<td>2.36</td>
<td>2.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dominant</td>
<td>1.70</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>D7S2212</td>
<td>Recessive</td>
<td>1.55</td>
<td>2.25</td>
</tr>
<tr>
<td>Median age of PC onset 56-72 years (214)</td>
<td>D6S1281</td>
<td>Dominant affected only</td>
<td>3.42</td>
<td>3.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dominant</td>
<td>2.52</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>D7S2212</td>
<td>Recessive</td>
<td>1.68</td>
<td>2.41</td>
</tr>
<tr>
<td>≥5 sampled affected (26)</td>
<td>D2S1391</td>
<td>Dominant</td>
<td>2.63</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>D8S1119</td>
<td>Recessive</td>
<td>2.01</td>
<td>2.01</td>
</tr>
<tr>
<td></td>
<td>D10S1432</td>
<td>Dominant</td>
<td>1.93</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td>D13S285</td>
<td>Recessive</td>
<td>2.21</td>
<td>2.21</td>
</tr>
</tbody>
</table>

Over 800,000 genotypes completed


Summary of Approximately 15 Individual Prostate Cancer Genome Wide Scans

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**

LOD scores

LOD = 1.86

LOD = 3.30

Fig 1. LOD in all 1,233 families
Combined Genome-Wide Screen Among 269 Families with ≥ 5 Affecteds

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

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

\[ \text{Lod} = 3.57 \text{ at 22q12} \]

ICPCG Genome-wide Scan
Chromosome 22 Linkage Results 1139 Families

- Overall
- Affecteds ≤ 2
- Affecteds = 3
- Affecteds = 4
- Affecteds ≥ 5
- \( \text{AAO} \leq 65 \)
- \( \text{AAO} > 65 \)
- Caucasians
- African Americans
- X-linked
- Not X-linked
- HPCa
- Non-HPCa

1 Lod Support Interval is 36-49 cM

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

\( \text{Lod} = \text{LOD} = 3.20 \)
\( \text{LOD} = 1.86 \)
Refining Region of Interest

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

Position in mega base-pair on chr 22, (Bld.36)
**Peak Region of Linkage**

**a)**
- Mayo, n=24
- PROGRESS, n=18
- Combined, n=42
- ICPCG, 3-recomb. interval, n=54
- Utah, n=14

**Position in mega base-pair on chromosome 22, (Bld.36)**

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

---

**Candidate Genes**

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

- ICPCG (incl. Utah)
- P+M+U All fam
- P+M All fam
- Utah
- P+M Fam with >5 cases

Position in Mega Base-pair on chromosome 22, (Bld.36)
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 \( \geq 3 \) men with aggressive disease (\( \geq 2 \) 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 \( \geq 7 \) or poorly differentiated grade
3. Prostate specific antigen at diagnosis \( \geq 20 \) 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**

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Subset</th>
<th>Position of max. LOD (cM)</th>
<th>Dm-LOD</th>
<th>Rec-LOD</th>
<th>IC-LOD&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Flanking markers (cM)</th>
<th>Marker (cM)</th>
<th>Marker (cM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>No. aff. ≥5</td>
<td>167.9</td>
<td>0.41</td>
<td>1.87</td>
<td>2.10</td>
<td>D2S1833 (62.4)</td>
<td>D2S1778 (172.9)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>HPC = No</td>
<td>69.2</td>
<td>1.51</td>
<td>1.47</td>
<td>2.06</td>
<td>D5S2680 (66.2)</td>
<td>GATA13B05 (75.9)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Dx age ≤ 58</td>
<td>124.8</td>
<td>1.75</td>
<td>2.16</td>
<td>1.42</td>
<td>D6S474 (117.6)</td>
<td>D6S410 (127.7)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>HPC = no</td>
<td>61.4</td>
<td>1.18</td>
<td>2.04</td>
<td>1.20</td>
<td>D6S1019 (93.4)</td>
<td>D6S1017 (62.8)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>No. aff. ≥ 5</td>
<td>7.4</td>
<td>3.16</td>
<td>0.97</td>
<td>1.80</td>
<td>D7S5058 (7.4)</td>
<td>D7S513 (17.0)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Dx age &lt; 65</td>
<td>46.2</td>
<td>0.63</td>
<td>1.47</td>
<td>2.25</td>
<td>D12S373 (2.7)</td>
<td>D12S1064 (48.0)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>No. aff. ≥ 5</td>
<td>103.6</td>
<td>2.07</td>
<td>0.65</td>
<td>0.96</td>
<td>D13S895 (47.9)</td>
<td>D13S285 (109.5)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>M to M = no&lt;sup&gt;6&lt;/sup&gt;</td>
<td>26.6</td>
<td>2.81</td>
<td>0.66</td>
<td>4.30</td>
<td>ATTC091 (26.4)</td>
<td>D8S504 (52.7)</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Dx age &lt; 65</td>
<td>43.9</td>
<td>0.78</td>
<td>2.77</td>
<td>2.06 (45.8)</td>
<td>D22S693 (33.7)</td>
<td>D22S699 (48.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dx age (69–70)</td>
<td>15.8</td>
<td>2.32</td>
<td>1.02</td>
<td>1.33</td>
<td>ATT019 (35.6)</td>
<td>D22S699 (28.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M to M = yes</td>
<td>15.6</td>
<td>3.75</td>
<td>1.79</td>
<td>2.05 (11.4)</td>
<td>ATT019 (35.6)</td>
<td>D22S699 (56.0)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Suggestive of X linkage.

<sup>b</sup>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

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

Johannesson et al., 2006, Prostate 15: 732-43
### Parametric Multipoint Analysis of Chromosome 11

![Graph showing HLOD values for different regions of chromosome 11.](image)

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

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

<table>
<thead>
<tr>
<th>band</th>
<th>Marker</th>
<th>Mbp</th>
<th>cM†</th>
<th>HLOD‡</th>
<th>α††</th>
<th>K&amp;C p-value‡‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>11p13</td>
<td>D11S1392</td>
<td>34.60</td>
<td>43.16</td>
<td>0.93</td>
<td>0.76</td>
<td>0.04</td>
</tr>
<tr>
<td>11p11.2</td>
<td>D11S1993</td>
<td>43.57</td>
<td>54.09</td>
<td>1.26</td>
<td>0.72</td>
<td>0.03</td>
</tr>
<tr>
<td>11p11.2</td>
<td>D11S190</td>
<td>44.98</td>
<td>54.50</td>
<td>3.10</td>
<td>1.00</td>
<td>0.004</td>
</tr>
<tr>
<td>11p11.2</td>
<td>D11S1395</td>
<td>51.23</td>
<td>56.33</td>
<td>3.17</td>
<td>1.00</td>
<td>0.005</td>
</tr>
<tr>
<td>Centromere</td>
<td>D11S1313</td>
<td>55.99</td>
<td>57.74</td>
<td>3.20</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>11q12.1</td>
<td>D11S4202</td>
<td>58.11</td>
<td>58.36</td>
<td>3.19</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>11q12.1</td>
<td>D11S1985</td>
<td>58.25</td>
<td>58.40</td>
<td>3.19</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>11q12.1</td>
<td>D11S4075</td>
<td>59.26</td>
<td>59.09</td>
<td>3.19</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>11q12.1</td>
<td>D11S1335</td>
<td>59.29</td>
<td>59.11</td>
<td>3.19</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>11q12.1</td>
<td>D11S2006</td>
<td>59.47</td>
<td>59.24</td>
<td>3.19</td>
<td>1.00</td>
<td>0.007</td>
</tr>
<tr>
<td>11q12.2</td>
<td>D11S4191</td>
<td>59.76</td>
<td>60.09</td>
<td>3.14</td>
<td>1.00</td>
<td>0.008</td>
</tr>
<tr>
<td>11q12.2</td>
<td>D11S1765</td>
<td>60.53</td>
<td>61.78</td>
<td>1.64</td>
<td>0.74</td>
<td>0.01</td>
</tr>
<tr>
<td>11q12.3</td>
<td>D11S4076</td>
<td>61.11</td>
<td>62.62</td>
<td>1.68</td>
<td>0.74</td>
<td>0.01</td>
</tr>
<tr>
<td>11q13.1</td>
<td>AAT268</td>
<td>62.82</td>
<td>64.60</td>
<td>1.70</td>
<td>0.73</td>
<td>0.02</td>
</tr>
<tr>
<td>11q13.2</td>
<td>D11S1883</td>
<td>63.12</td>
<td>64.97</td>
<td>1.63</td>
<td>0.73</td>
<td>0.02</td>
</tr>
<tr>
<td>11q13.2</td>
<td>D11S913</td>
<td>65.68</td>
<td>67.40</td>
<td>1.24</td>
<td>0.73</td>
<td>0.06</td>
</tr>
<tr>
<td>11q13.2</td>
<td>D11S1889</td>
<td>67.06</td>
<td>69.28</td>
<td>0.36</td>
<td>0.43</td>
<td>0.14</td>
</tr>
<tr>
<td>11q13.3</td>
<td>D11S987</td>
<td>67.65</td>
<td>69.94</td>
<td>0.23</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>11q13.4</td>
<td>D11S4136</td>
<td>69.31</td>
<td>71.52</td>
<td>0.16</td>
<td>0.26</td>
<td>0.20</td>
</tr>
<tr>
<td>11q13.4</td>
<td>D11S4162</td>
<td>70.64</td>
<td>72.75</td>
<td>0.19</td>
<td>0.30</td>
<td>0.20</td>
</tr>
<tr>
<td>11q13.4</td>
<td>D11S2371</td>
<td>73.18</td>
<td>76.13</td>
<td>0.39</td>
<td>0.40</td>
<td>0.20</td>
</tr>
</tbody>
</table>

*Johannesson et al., 2006, Prostate 15: 732-43*
Candidate Region on Chromosome 11 for Kidney-Prostate Cancer

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

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

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?

Founder Chromosome

Many Generations

Today

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