

Genetic Risk Factors for Childhood Leukemia

GETA Symposium
3 March, 2009

Anand Chokkalingam, Ph.D.
School of Public Health,
University of California, Berkeley
anandc@berkeley.edu

Outline

- Background
- Review from literature
- Northern California Childhood Leukemia Study (NCCCLS) experience
- Challenges
- Future directions



Genetic susceptibility

- Q: What is genetic susceptibility?
- A: Heritable factors that increase risk of a given disease
- Usually one or more genes, or gene variations
- May work in concert with
 - Other genetic factors, AND/OR
 - Environmental and lifestyle factors
- Degree of involvement of other factors depends on **penetrance**



Penetrance

- Penetrance

- High: rare, but high risk (e.g., BRCA1, RR~5)
 - Major part of short causal pathway to disease



- Low: common, but low risk (RRs of ~1.3-1.8)
 - Minor part of long causal pathway to disease



- Low-penetrance genetic factors likely to comprise the bulk of inherited cancer risk



Background

Review

NCCLS

Challenges

Future
Directions

Rationale for genetic susceptibility to CL

- Early age of onset
- Risk in twins
 - Mostly intraplacental metastasis, not highly penetrant risk allele
 - Suggests low penetrance susceptibility alleles



3 March, 2009

5

Review

Background

Review

NCCLS

Challenges

Future
Directions

Candidate pathways in published reports

- Folate metabolism
- Xenobiotic (exogenous chemical) transport and metabolism
- Immune function
- DNA repair



3 March, 2009

Primary reports of gene main effects: ALL

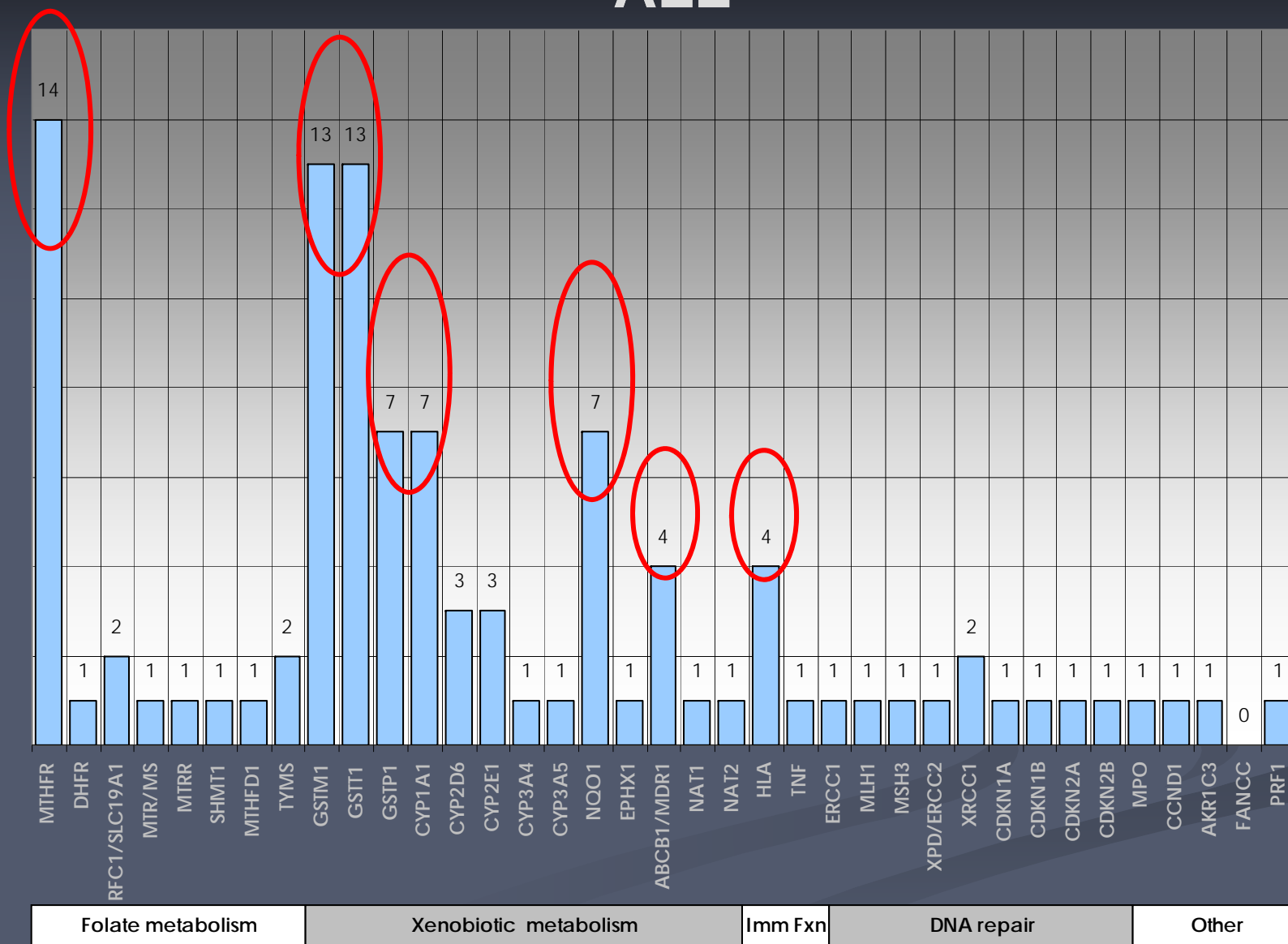
Background

Review

NCCLS

Challenges

Future
Directions



3 March, 2009

8

Folate metabolism & ALL

- Folate

- Essential micronutrient, modulates balance between accuracy of DNA synthesis and DNA methylation
- Deficiency can induce chromosomal damage and fragile chromosomal sites → carcinogenesis
- Maternal supplementation during pregnancy may reduce risk

- MTHFR, 2 loss-of-function variants: 14 reports

- 677C>T: null effect or modest risk reduction (OR~0.9)
- 1298A>C: null effect



Xenobiotic metabolism & ALL

- To do harm, exogenous chemicals must
 - enter cells
 - Membrane transporters (e.g., MDR1)
 - be metabolized into harmful species
 - Phase 1, bioactivation enzymes (e.g., CYPs)
 - Phase 2, detoxification enzymes (e.g., GSTs, NQO1)
- Transporters
 - MDR1: 4 reports
 - 3435C>T: null risk
- Phase 1, bioactivation
 - CYP1A1: 7 repots
 - 6235T>C: inconsistent



Xenobiotic metabolism & ALL

- Phase 2, detoxification
 - GSTM1 (detoxifies PAHs): 13 reports
 - deletion: null to modestly increased risk
 - GSTT1 (detoxifies epoxides and halomethanes): 13 reports
 - deletion: null risk
 - GSTP1: 7 reports
 - I105V: null risk
 - NQO1 (anti-oxidant, detoxifies quinones): 7 reports
 - 609C>T: null risk
 - 465C>T: null risk (2 reports)



Primary reports of gene main effects: AML

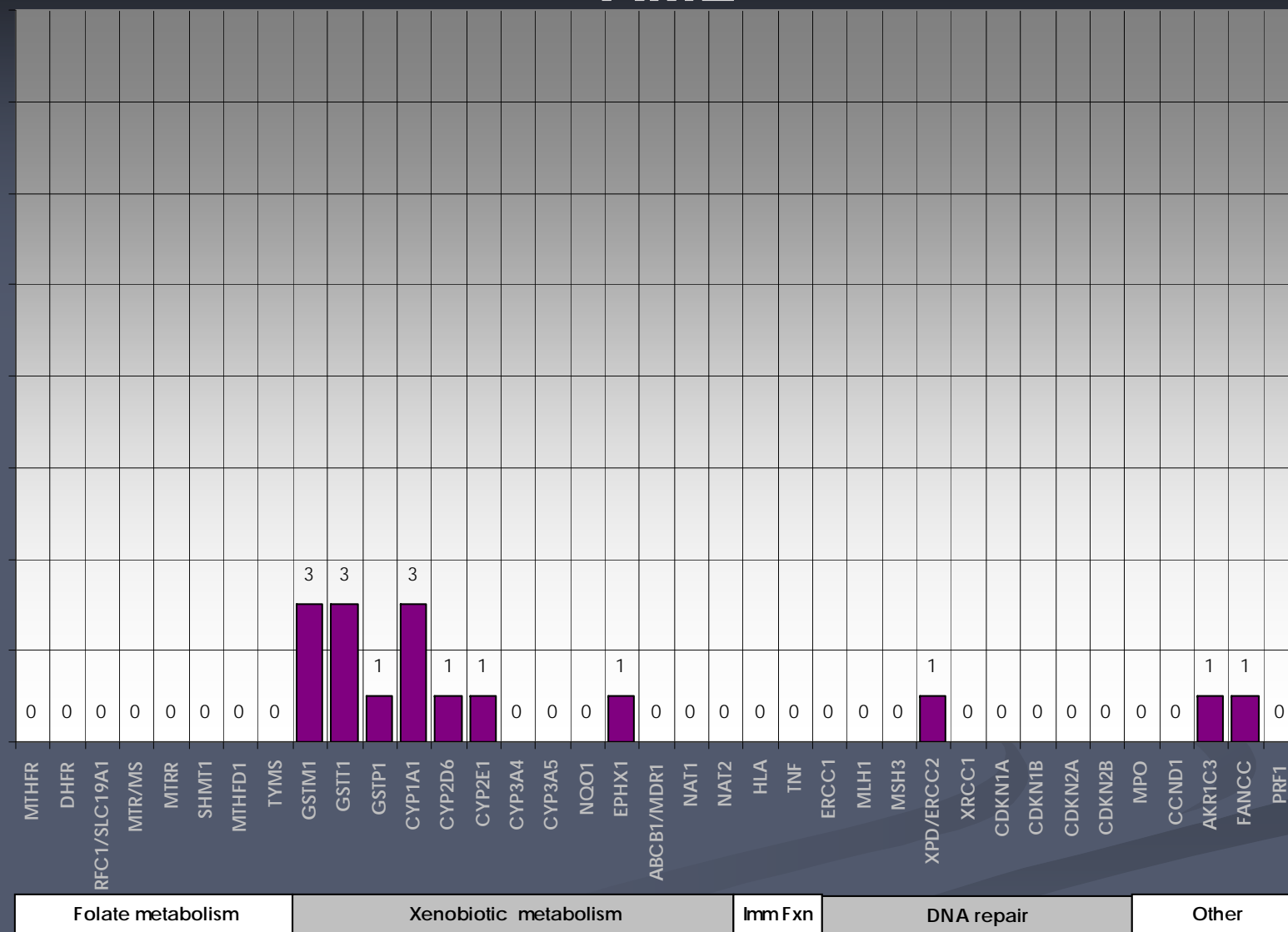
Background

Review

NCCLS

Challenges

Future
Directions



3 March, 2009

12

Background

Review

NCCLS

Challenges

Future
Directions

Evaluating the evidence

- HuGENet – Human Genome Epidemiology Network
- HuGENet Encyclopedia: synopsis of evidence for genetic associations of complex disease
- CL as one of several prototype encyclopedia entries
- Venice meeting (2006): to develop criteria for rapid evaluation of evidence to facilitate encyclopedia effort
 - 3 criteria:
 - Amount of evidence
 - Replication
 - Protection from bias
 - Letter grades (A, B, C) for each – AAA is ideal

Published by Oxford University Press on behalf of the International Epidemiological Association
© The Author 2007; all rights reserved.

International Journal of Epidemiology 2007;1–11
doi:10.1093/ije/dym159

Assessment of cumulative evidence on genetic associations: interim guidelines

John P A Ioannidis,^{1–3*} Paolo Boffetta,⁴ Julian Little,⁵ Thomas R O'Brien,⁶ Andre G Uitterlinden,⁷ Paolo Vineis,⁸ David J Balding,⁸ Anand Chokkalingam,⁹ Siobhan M Dolan,¹⁰ W Dana Flanders,¹¹ Julian P T Higgins,¹² Mark I McCarthy,^{13,14} David H McDermott,¹⁵ Grier P Page,¹⁶ Timothy R Rebbeck,¹⁷ Daniela Seminara¹⁸ and Muin J Khoury¹⁹



3 March, 2008

13

Venice criteria evaluation for ALL

Pilot Results

- Result of preliminary review
 - None have reached A status in any criterion
 - Only MTHFR and GSTM1 rank BBB
 - All else have a C in at least one criterion
- Evaluation of criteria in progress
- Next steps:
 - Refine criteria
 - Develop systems for
 - Consistent assignment and adjudication of grading
 - Updating as new evidence is published



Summary

- Few genes have been studied to date for ALL, even fewer for AML
- Many high-probability candidates remain unexamined or unreplicated
- Entire candidate pathways with very strong biological plausibility remain poorly studied (e.g., immune function, DNA repair)
- Reports to date do not ensure good coverage of variation within a gene



Background

Review

NCCLS

Challenges

Future
Directions

HapMap Project

Publicly available SNP data



The screenshot shows the International HapMap Project website. At the top, there is a header with the project logo (a world map with a DNA double helix) and the title "International HapMap Project". Below the header, there is a navigation bar with links: Home | About the Project | Data | Publications | Tutorial.

The main content area includes a section titled "Instructions" with a search box and a "Search" button. Below the search box, there are examples of search terms: "Chr20", "Chr9:660,000..760,000", "SNP:rs6870660", "NM_153254", "BRCA2", "ENM010". There are also links for "[Help]" and "[Reset]".

Below the search box, there is a section titled "Search" with a "Help links:" label and a row of buttons: "- LD -", "- tagSNPs -", "- Phased Haplotype -", "- Genotype data -", "- Frequency data -", and "- Symbols and colours used -".

Below the search box, there is a section titled "Landmark or Region :" with a text input field and a "Search" button. To the right of this section, there is a "Reports & Analysis :" section with a dropdown menu showing "Annotate LD Plot" and buttons for "Configure..." and "Go".

Below the search box, there is a section titled "Data Source" with a dropdown menu showing "HapMap Data Rel#19/phaseII Oct05, on NCBI B34 assembly, dbSNP b124".

Below the search box, there is a section titled "Population descriptors:" with text describing the populations: "YRI: Yoruba in Ibadan, Nigeria, JPT: Japanese in Tokyo, Japan, CHB: Han Chinese in Beijing, China, CEU: CEPH (Utah residents with ancestry from northern and western Europe)".

Below the search box, there is a section titled "Overview" and "Details" with links for "Overview" and "Details".

At the bottom of the page, there is a note: "For performing in depth LD and Haplotype analysis of genotype data install Haploview in your local machine. Haploview (ver3.12) is now available for download."

- Data on linkage between SNPs in different populations

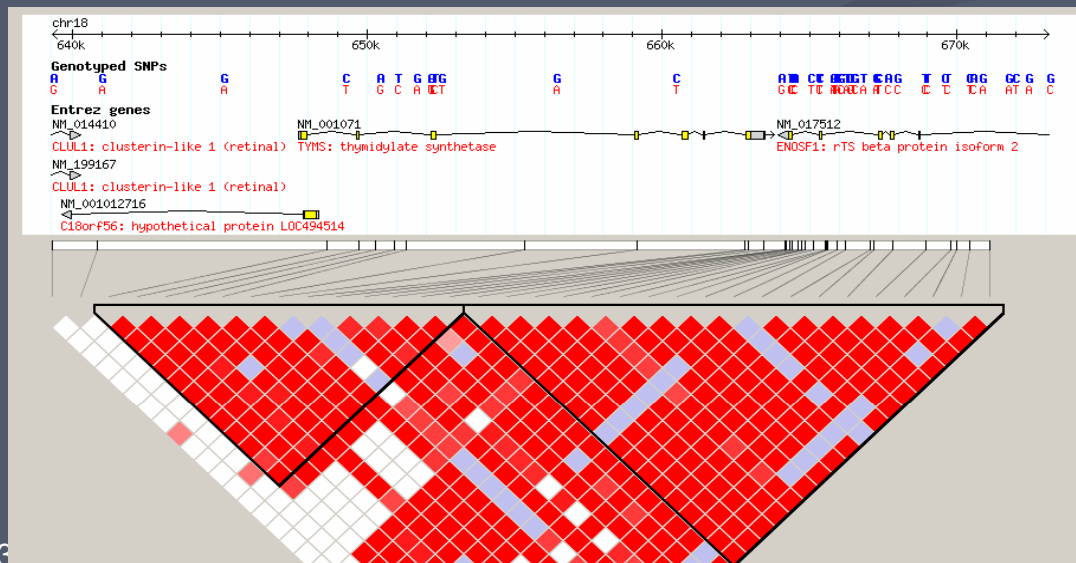


3 March, 2009

16

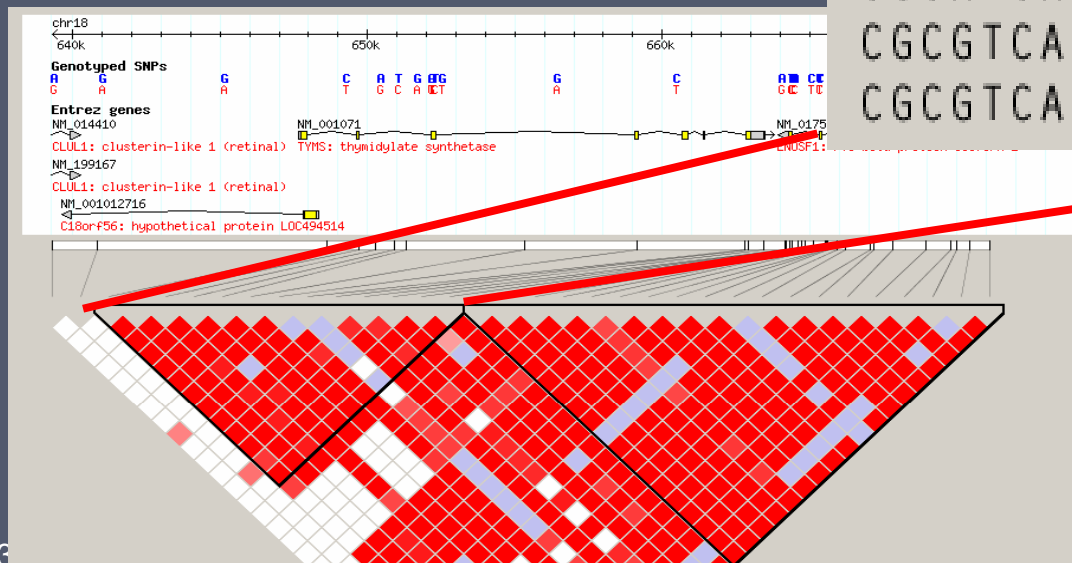
Haplotype-based analysis

- Uses HapMap and similar data
- Permits:
 - Maximal coverage of variation within genes
 - Minimum number of SNPs



Haplotype-based analysis

- Uses HapMap and similar data
- Permits:
 - Maximal coverage of genes
 - Minimum number of



Block 1

SNP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
TATGCTGCGGCCGC																					.508
CGCATCATACTTAT																					.200
CGCGTCACGGCCGC																					.175
CGCGTCATACTTAT																					.067



NCCLS experience

Background

Review

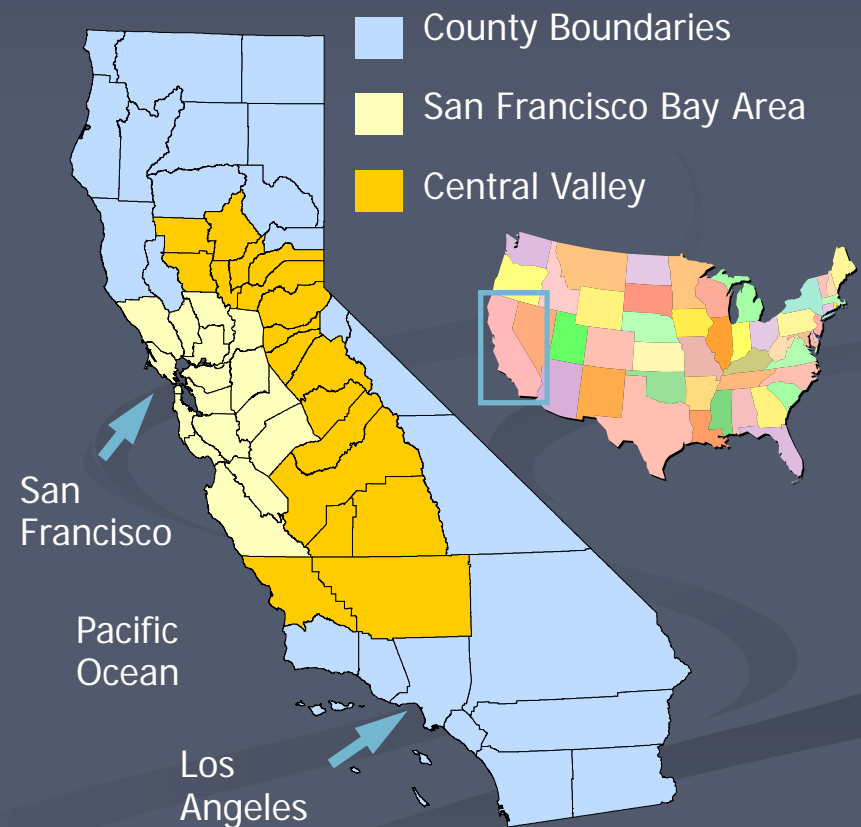
NCCLS

Challenges

Future
Directions

Northern California Childhood Leukemia Study 1995-present

- Population-based case-control study
- Incident cases ascertained from 9 pediatric hospitals in N. & C. California
- Controls individually matched (DOB, sex, Hispanic status, and maternal race)
- 42% Hispanic



3 March, 2009

20

Background

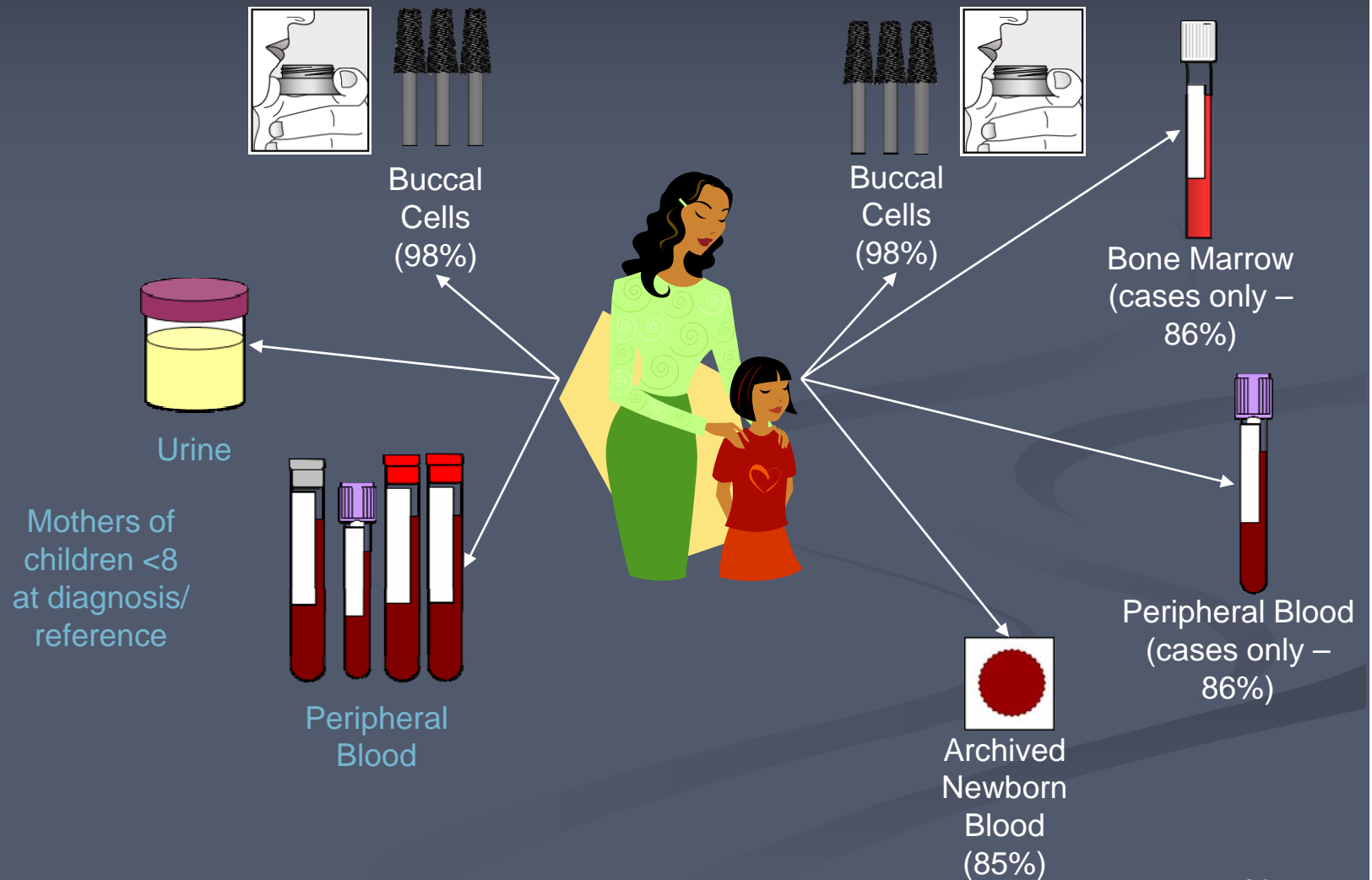
Review

NCCLS

Challenges

Future
Directions

NCCLS Biospecimen collection



3 March, 2009

21

Background

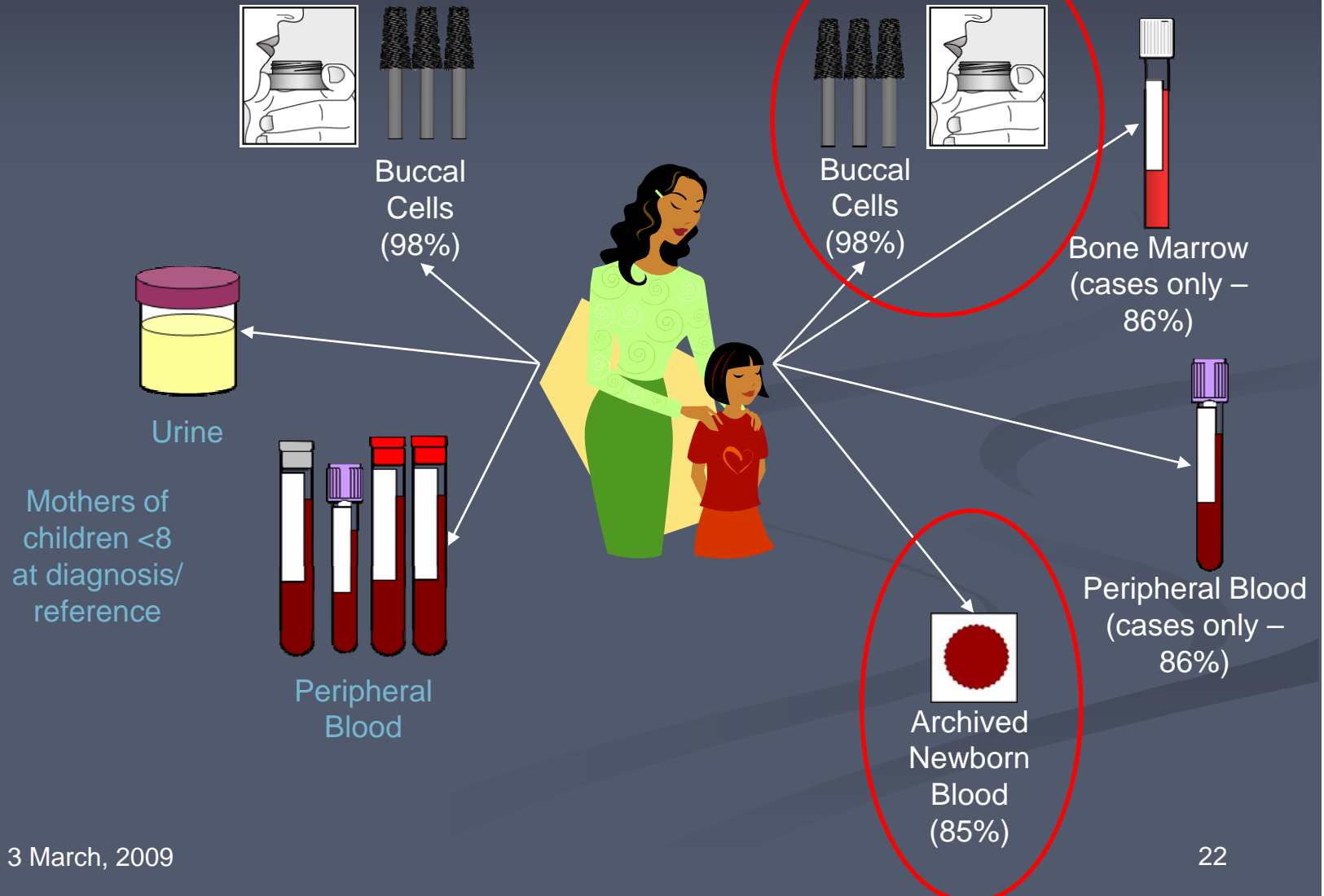
Review

NCCLS

Challenges

Future
Directions

NCCLS Biospecimen collection



3 March, 2009

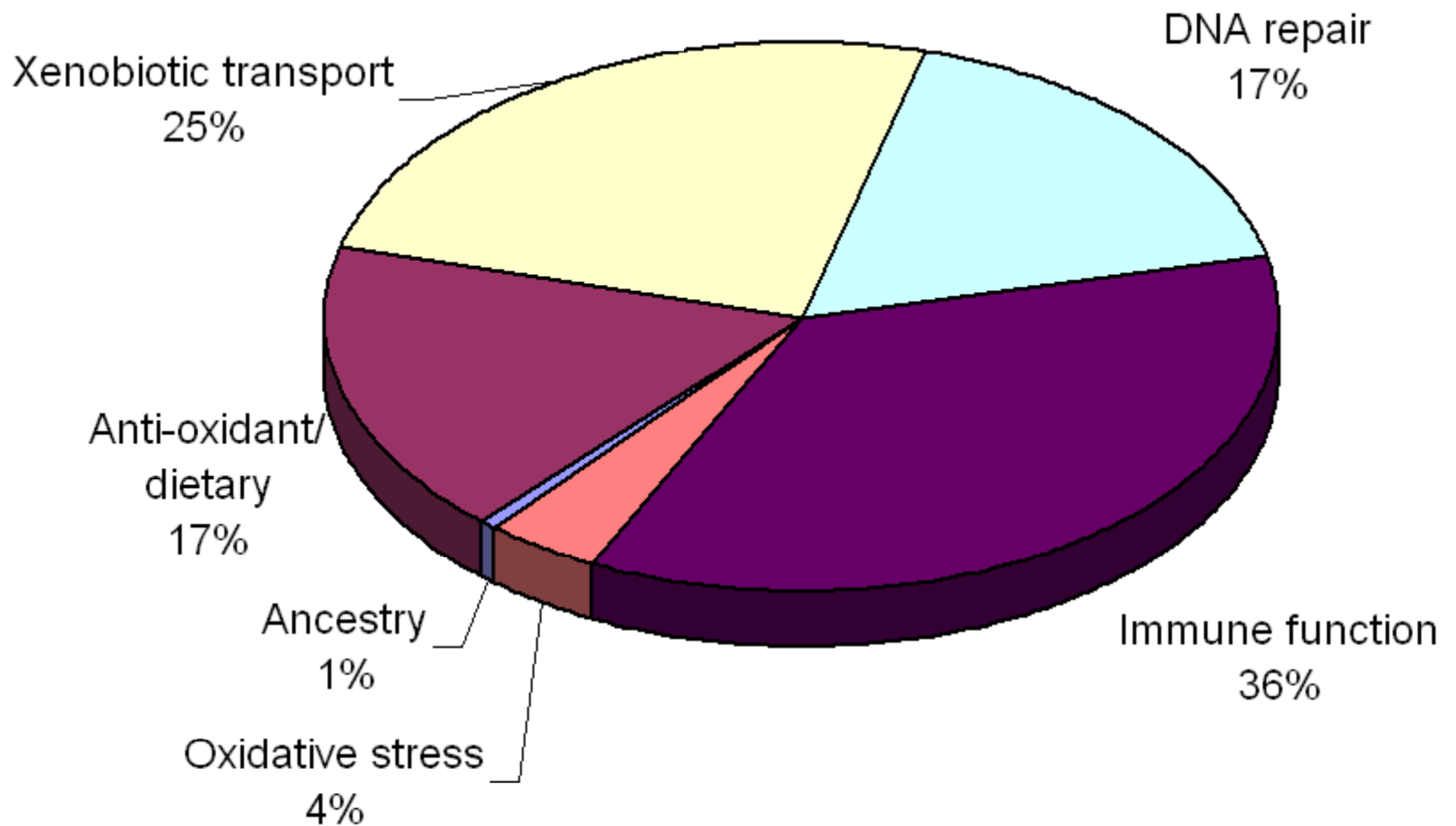
22

Large-scale Genotyping

- **Objective:** to comprehensively examine ~200 candidate genes in subset of available children (464 cases, 464 controls)
- Custom Illumina 1536-plex
 - 183 genes
 - Haplotype tagging SNPs
 - Literature SNPs
 - Ancestry Informative Markers
 - Adjust for genetic ancestry



Candidate Pathways



Background

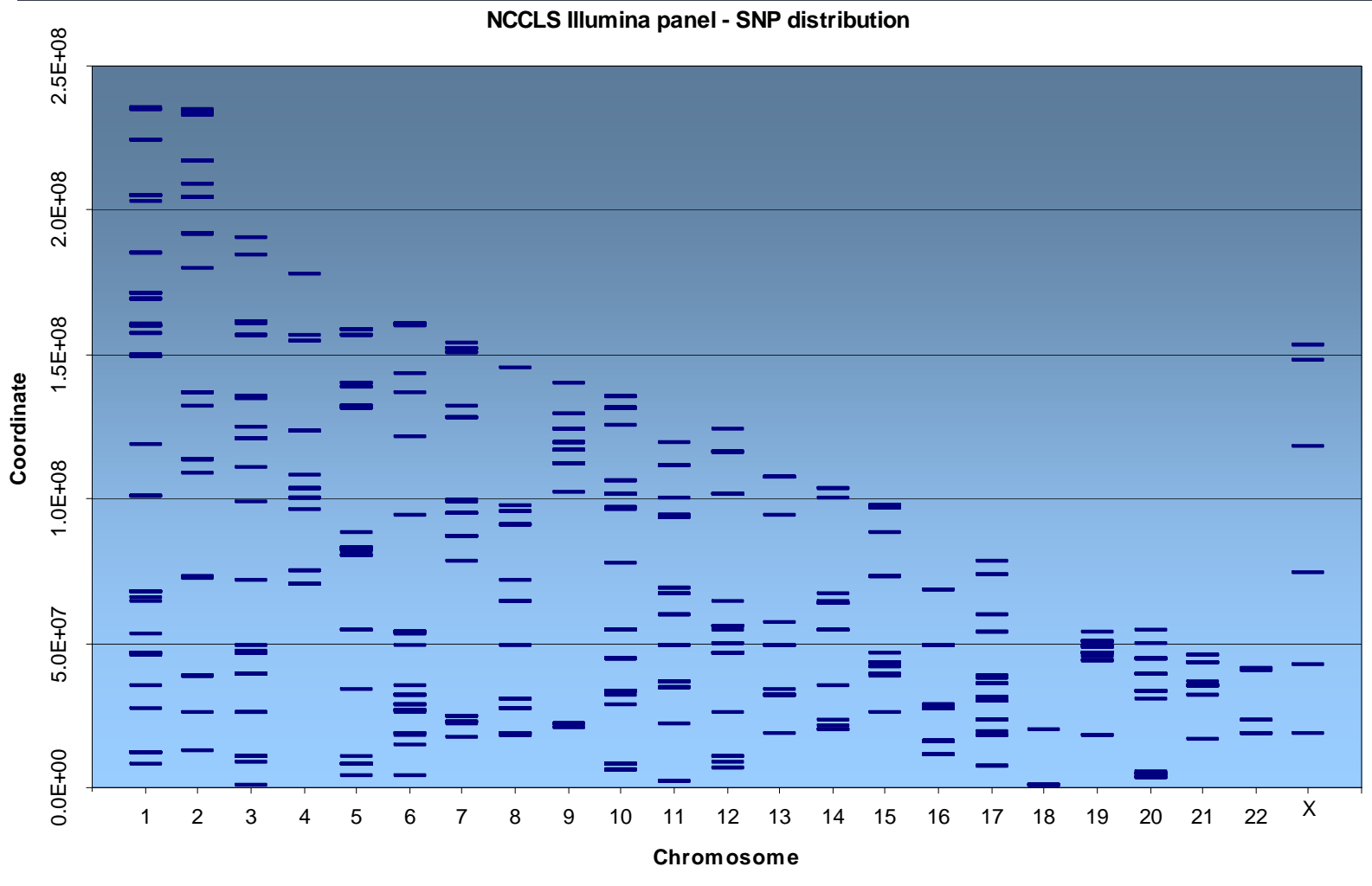
Review

NCCLS

Challenges

Future
Directions

NCCLS panel coverage



3 March, 2009

25

NCCLS genotyping

Current status – Preliminary analyses

- Adaptive immunity and ALL (208 SNPs in 29 genes)
 - 19 SNPs with nominal $p < 0.05$
 - Single *IL12A* SNP ($p < 0.00001$)
 - Haplotype sliding window analyses: regions in *CD28*, *FCGR2A*, *GATA3*, *IL2RA*, *IL12A*, *STAT4*, and *STAT6*
- Xenobiotic transport and metabolism and ALL (251 SNPs in 42 genes)
 - Individual SNP associations for *ABCB1*, *ABCC1*, and *CYP2C8* genes.
 - Haplotype sliding window analyses: regions in *ABCC1*, *ABCB1*, *CYP1A1*, *CYP1A2*, *CYP2B6*, *CYP2C8*, *IDH1*, *UGT1A1*, *UGT1A7*, and *UGT1A9*.
 - Potential interaction: *CYP2C8* risk haplotype and self-reported exposure to paints and solvents



Challenges to genetic susceptibility research in childhood leukemia

Statistical power and sample size issues

Maternal genes

Replication

Publication Bias

Background

Review

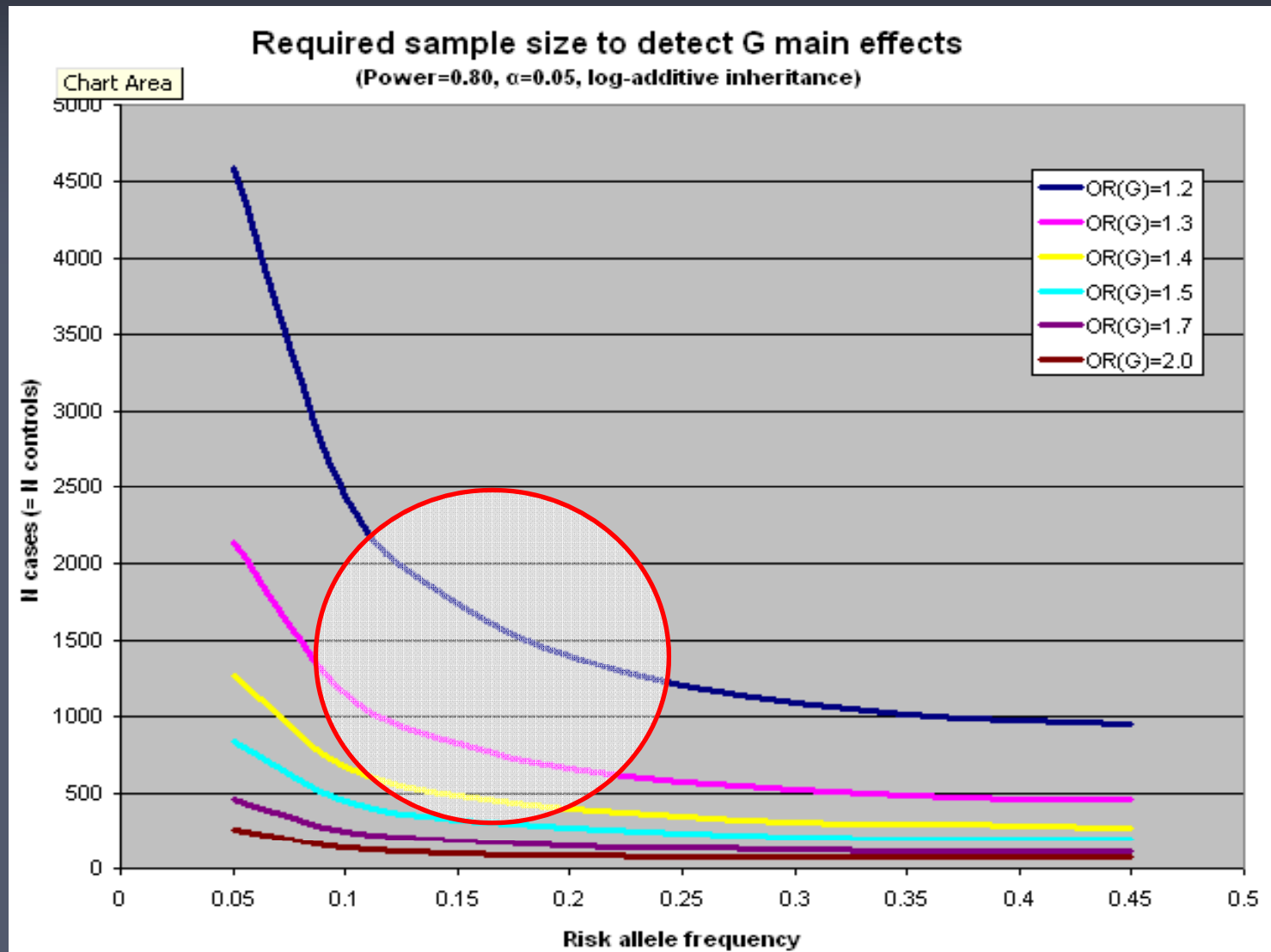
NCCLS

Challenges

Future
Directions



Statistical power

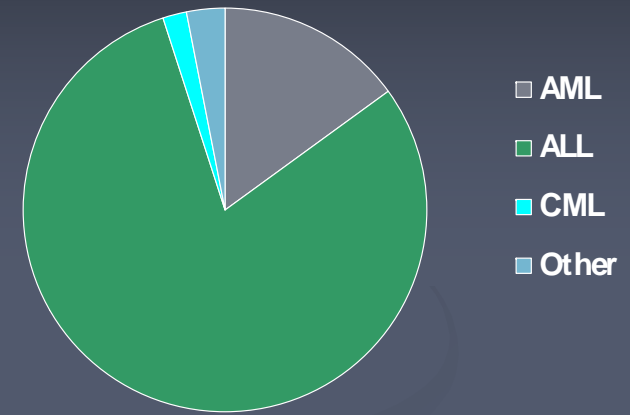


3 March, 2009

28

Sample size

- Sample sizes limited by
 - Low incidence
 - 4.5/100,00 person-years
 - Disease heterogeneity
 - Subtype groups
 - “Lumping vs. splitting”
 - Availability of high-quality DNA
- Published genetic studies typically 100-300 cases
 - Larger studies expected



Interactions

- Small effect sizes (OR~1.2-1.5) point to effects of GxE interactions
- Critical to understanding susceptibility
- Requires
 - High-quality exposure data
 - Even larger sample sizes for sufficient power



Background

Review

NCCLS

Challenges

Future
Directions

Maternal genes



3 March, 2009

31

Other major challenges

- Replication of results
 - Multiple studies
 - Multiple populations
- Publication bias
 - When null results are not presented or published



Future directions

CLIC

- Childhood Leukemia International Consortium
 - www.clic.berkeley.edu
 - To date: 14 case-control studies in 10 countries
 - Over 9,000 cases
- Purpose
 - Replication of findings
 - Coordinated publication (address publication bias)
 - Data pooling (improve statistical power)
 - Collaborative research



GWAS

- Genome-Wide Association Studies
 - Allows exploration of genome beyond candidate genes
 - 300K-1M variants across the genome
- Issues
 - Sample size
 - Replication strategy
 - Technical requirements
 - DNA quality (unamplified, genomic DNA)
 - DNA quantity (~1 microgram)
 - Cost



Background

Review

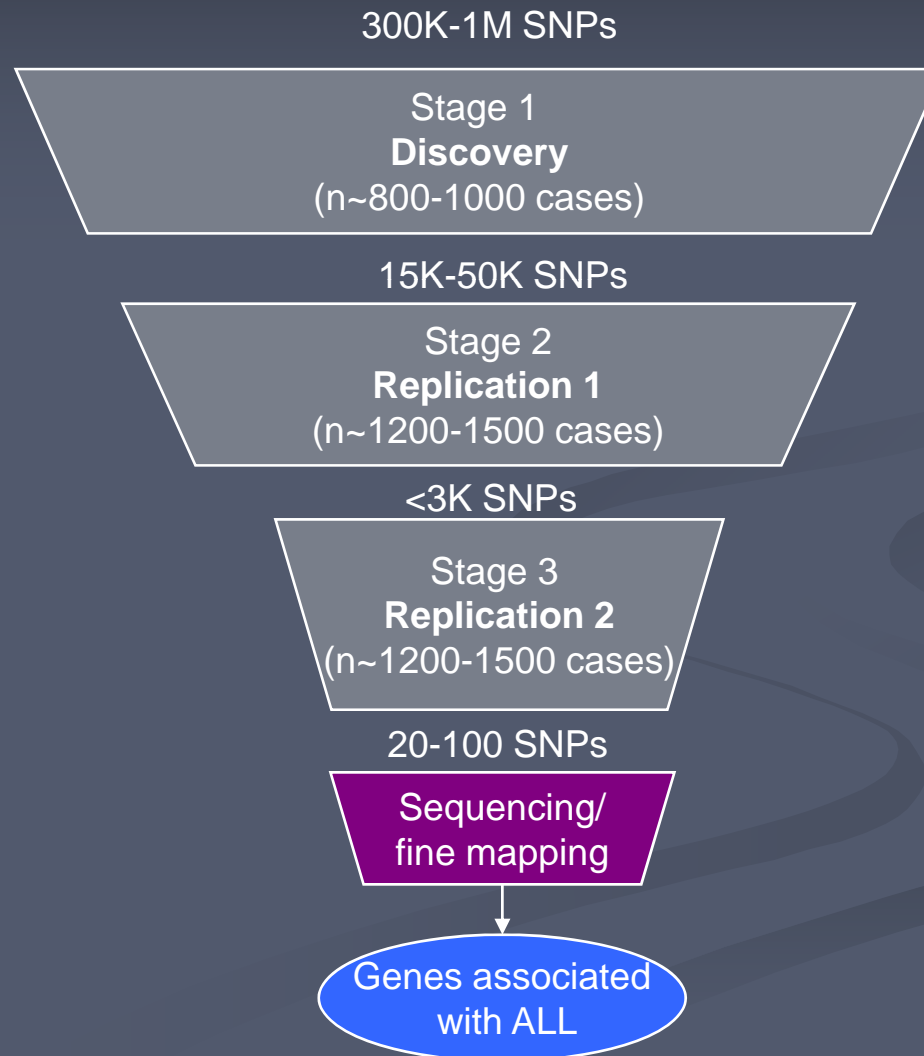
NCCLS

Challenges

Future
Directions

GWAS:

Tiered replication approach



3 March, 2009

36

Acknowledgments

- UCB

- Patricia Buffler
- Catherine Metayer
- Kevin Urayama
- Ghislaine Scelo
- Jeffrey Chang
- Melinda Aldrich
- Lisa Barcellos
- Martyn Smith

- UCSF

- Joe Wiemels
- John Wiencke
- Helen Hansen

- CDC

- Muin Khoury

- NCI

- Stephen Chanock

- UKCCS

- Malcolm Taylor

