Genetic Research in Utah

Lisa Cannon Albright, PhD
Professor, Program Leader
Genetic Epidemiology
Department of Internal Medicine
University of Utah School of Medicine

George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, Utah

Huntsman Cancer Institute, Salt Lake City, Utah

Today's Outline

The Utah Genealogy Resource

Heritable Contribution to Parkinson's Disease

Utah Gene Discovery

Genealogies and Medicine

Today's Outline

The Utah Genealogy Resource

Heritable Contribution to Parkinson's Disease

Utah Gene Discovery

Genealogies and Medicine

Utah Genealogy - Utah Population Data Base (UPDB)

In the 1960's Mark Skolnick, PhD constructed the (first) computerized genealogy of the Parma Valley, Italy, using parish records; medical records were diffuse in this population

In the 1970's he built the Utah Genealogy, linked it to cancer and death data and began to study high-risk breast cancer pedigrees – the Utah Genealogy was supported by NIH and later donated to the University of Utah and became the UPDB

In the 1990's Mark co-founded Myriad Genetics. His University group, Genetic Epidemiology, collaborated with Myriad using Utah high-risk pedigrees and analysis methods to localize genes; we jointly discovered *BRCA1*, *BRCA2*, and *CDKN2A* as the first major cancer predisposition genes

The Utah Population

Adherence to proscriptions against coffee, tea, tobacco, alcohol high

Teachings encourage large families with strong ties, high educational attainment, and strict sexual mores

During the late nineteenth century, mean number of offspring per couple was > 5 during each decade and reached almost 9

Before 1890 polygamy was practiced among the pioneers, leading to families with multiple wives and dozens of children

Founding pioneers were largely unrelated; low/normal inbreeding

Receptive to research studies

Original Utah Genealogy Data

Mormons make up 75% of the state of Utah

- individual Mormons trace their ancestries as far as possible
- records collected in the Family History Library of the Church

The Utah Genealogy used 3-generation family genealogy sheets submitted by members of the LDS Church

Skolnick selected sheets containing at least one life event in Utah or on the pioneer trail (1840-1850); record linking accomplished during data entry

Original Utah genealogy included 1.6 million individuals in genealogies 6 - 7 generations deep

Genealogy now extended with Utah vital statistics data (e.g. trios from birth certificates)

- Pedigrees up to 16 generations

Records available in UPDB today

Record Type	Number of Records
Genealogy Records	1,625,707
US Utah 1880 - 1940 Census	2,300,087
Birth Certificates (1915-21, 1935-2014)	2,776,515
Marriage Certificates (1978-2010)	689,052
Fetal Deaths (1978-2012)	9,637
Social Security Death Index (Nationwide)	581,371
Inpatient Hospital Claims (1996-2006)	2,726,004
Ambulatory Surgery Utah (1996-2013)	4,912,396
Cancer Registry Utah (1966-2013)	324,350
Death Certificates (1904 – 2014)	847,028
Drivers License Utah	3,667,493
Inpatient hospital claims Utah (1996-2013)	4,677,782
Total Records	**24.7 million**

Utah Genealogy and Linked Phenotype Data

Genealogy of ~2.5 million Utah Mormon pioneers and descendants with linked phenotype

Other linked data

-University of Utah Health Sciences Center (UUHSC) Enterprise Data Resource

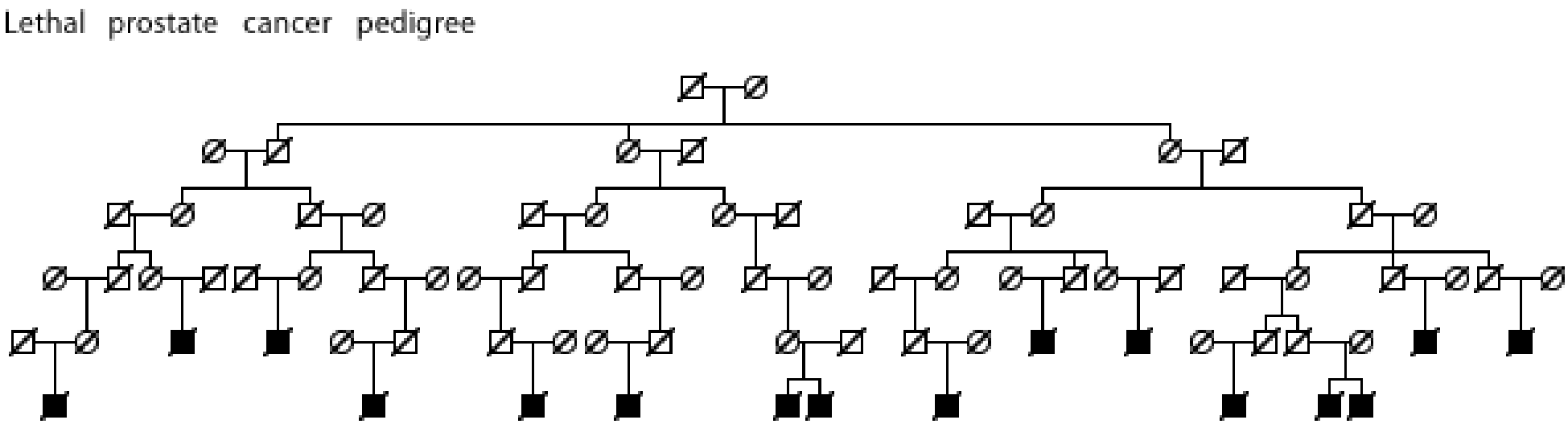
over 1.4 million patients

-Intermountain Healthcare EDW (largest health care provider in Utah)

over 3 million patients

http://www.hci.utah.edu/groups/ppr

Example Utah High-Risk Pedigree



Today's Outline

The Utah Genealogy Resource

Heritable Contribution to Parkinson's Disease

Utah Gene Discovery

Genealogies and Medicine

Identification of a Heritable Contribution to Parkinson's Disease with UPDB data

We use 3 different analyses to demonstrate the genetic contribution to Parkinson's Disease using the UPDB:

- Familial clustering (average relatedness) of patients
- Relative risk of Parkinson's Disease in relatives of patients
- Identification of pedigrees with an excess of patients

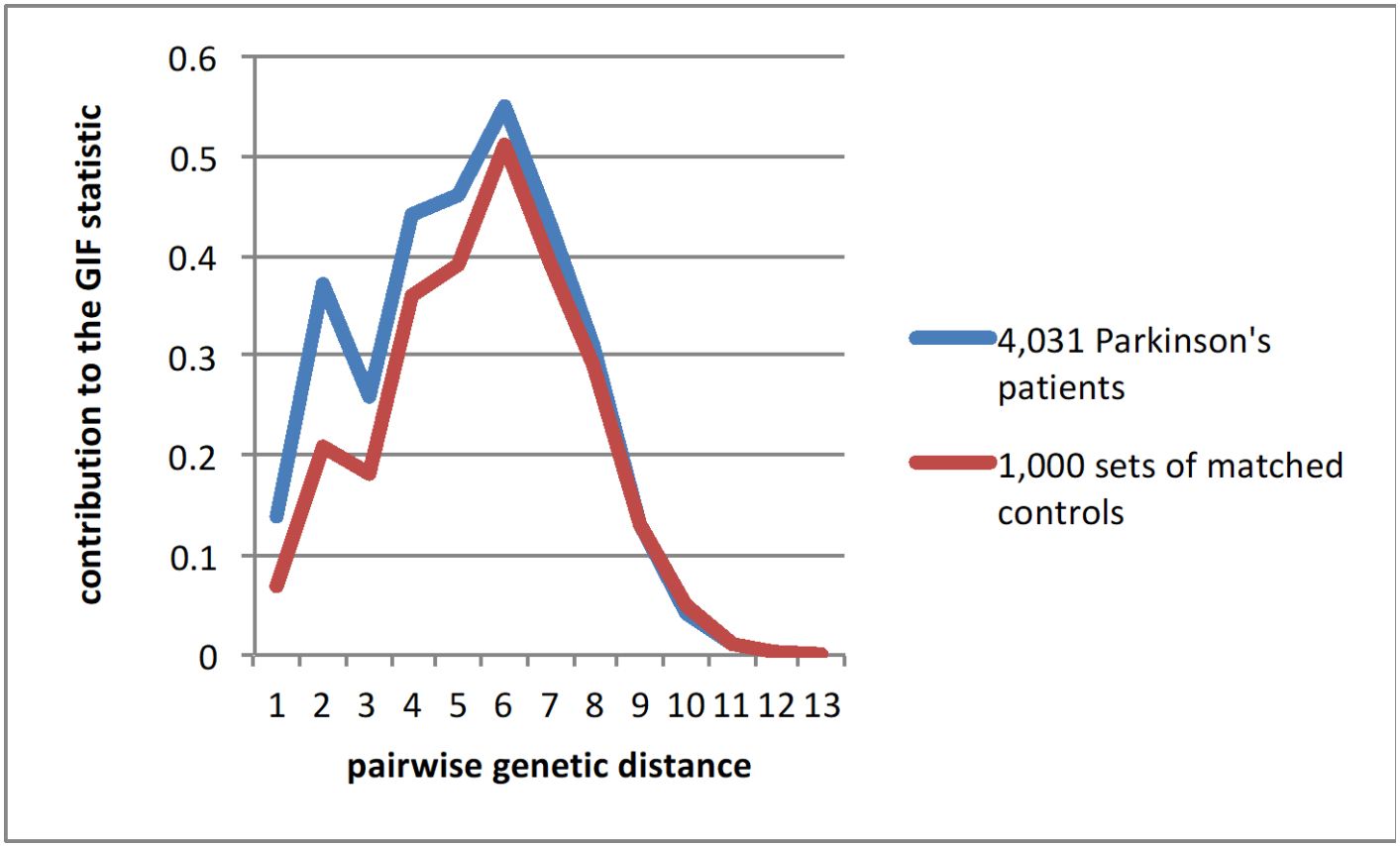
Utah Analysis of Familial Clustering

Concept: If Parkinson's has a familial/genetic component, Parkinson's patients should be more closely related to each other than people just like them in Utah

Method: Using a genealogy linked to Parkinson's Disease diagnosis we can define the genetic relationships between all the pairs of patients, and estimate the average "relatedness" of all individuals with Parkinson's.

The same measurement of relatedness on a set of matched controls estimates the average relatedness *expected* in the population

Utah Analysis of Familial Clustering of Parkinsons's Disease



Test for excess relatedness $p < 0.001$

Test for excess relatedness ignoring close relationships (1^{st} cousins and closer) $p = 0.001$

Relative Risk of Parkinson's in Relatives of Patients with Parkinson's

Concept: If Parkinson's has a heritable component, it should occur at higher-than-population rate among relatives of patients.

Method: Compare rate of Parkinson's Disease in relatives of cases to the population rate (estimated in the UPDB)

Interpretation: Excess first degree relative risks may be due to genes, environment, or both

Excess risks in more distant relatives strongly support a heritable component

Risks for Parkinsons' Disease in First- to Third- Degree Relatives

N = 4,031 patients with Parkinson's Disease with genealogy in the UPDB

18,127 first degree relatives	(parents, children, siblings)
40,546 second-degree relatives	(grandparent, grandchild, uncle, aunt, niece, nephew)
93,398 third-degree relatives	(cousin, great grandparent, greatgrandchild, great aunt/uncle/niece/nephew)

Parkinson's Diagnosis in:	Observed	Expected	Relative Risk	95% CI
First-degree relatives	273	150.4	1.82	1.61, 2.04
Second-degree relatives	325	225.9	1.44	1.29, 1.60
Third-degree relatives	804	727.8	1.10	1.03, 1.18

Pedigrees with a significant excess of Parkinson's Disease

Concept: if the descendants of a couple includes a higher number of Parkinson's patients than expected based on population rates, the pedigree is high-risk

Method: Identify patients in the UPDB and find all sets of Parkinson's patients descending from a pair of founders

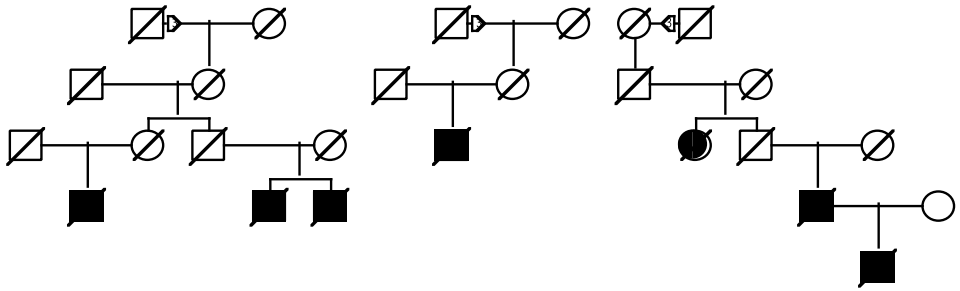
Compare observed and expected numbers of Parkinson's patients among descendants of the founders using rates of Parkinson's from UPDB

Pedigrees with significant excess of observed to expected patients are high-risk

Big Utah Difference: Typically not possible for other groups to identify "high-risk" pedigrees (as opposed to pedigrees with lots of related cases)

Utah Pedigree with an excess of Parkinson's Disease

Founder born in early 1800s in England, with 3 wives and 221 descendants in UPDB;
7 observed with Parkinson's, 1.8 patients expected; p=0.0003



There are hundreds of high risk Utah pedigrees

Today's Outline

The Utah Genealogy Resource

Heritable Contribution to Parkinson's Disease

Utah Gene Discovery

Genealogies and Medicine

Genetic Epidemiology High Risk Utah Pedigree Studies

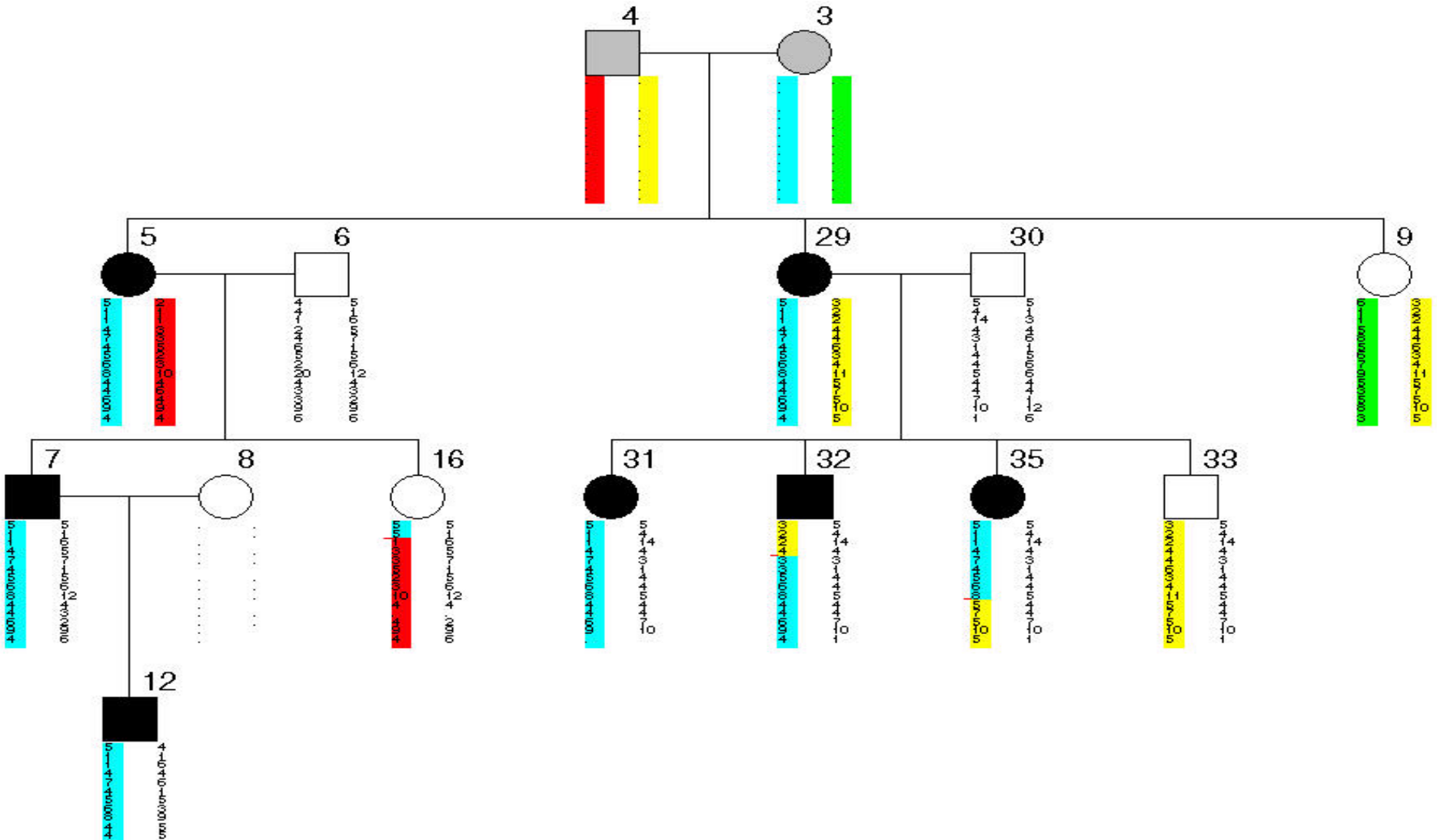
Disease		Pedigrees	DNAs
Breast Cancer		490	7536
Colon Cancer		272	4112
Celiac		319	2351
Melanoma		179	3075
Prostate Cancer		562	9737
Asthma		198	2494
Osteoporosis		323	1918
Depression		438	2939
Intracranial Aneurysm	151	721	
Pelvic Organ Prolapse		244	824
Rotator Cuff		127	544
Pancreas Cancer		107	180
Brain Cancer		20	50
Lung Cancer		20	60

Total 38,000+ DNAs

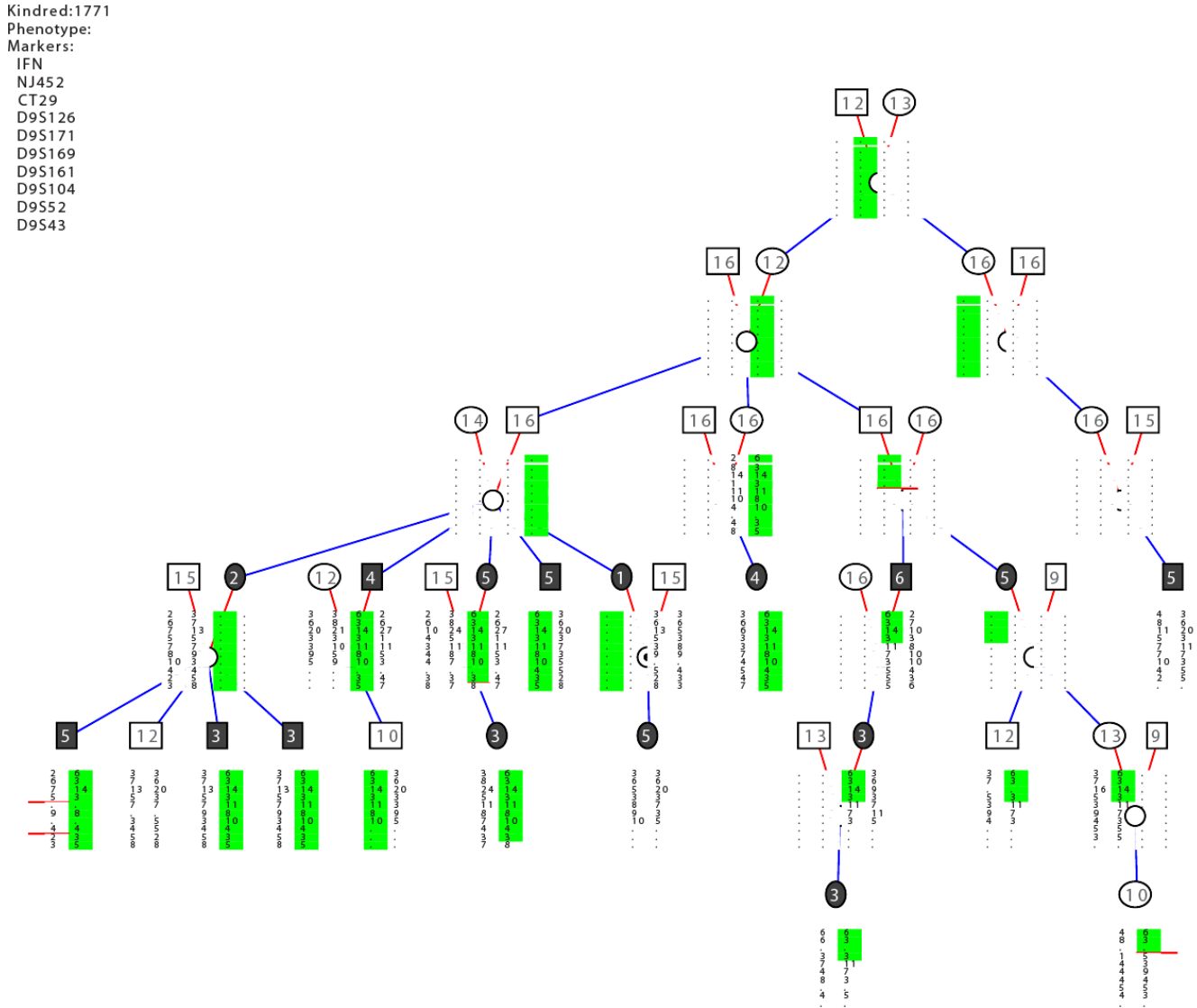
Utah Gene Discovery - Linkage Analysis in Extended Pedigrees

- Identify high-risk pedigrees
- Sample affected individuals and connecting relatives
- Genotype pedigree members
- Analyze co-segregation of genetic markers and phenotype to identify regions of chromosomes shared by cases
- Identify pedigrees exhibiting linkage evidence ($p<0.05$)
- Identify chromosomal region shared by cases in linked pedigrees

Analyzing Cosegregation of markers in Linked Pedigrees to Localize Gene



Melanoma Linkage - CDKN2A gene

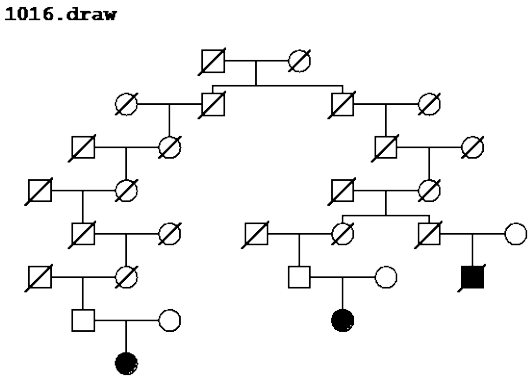


K1771 +3.17
14/16 cases share

New Utah Gene Discovery Approaches

Shared Genomic Segments

Search for chromosomal sharing Identical by Descent (IBD) in distantly related cases (~ linkage)

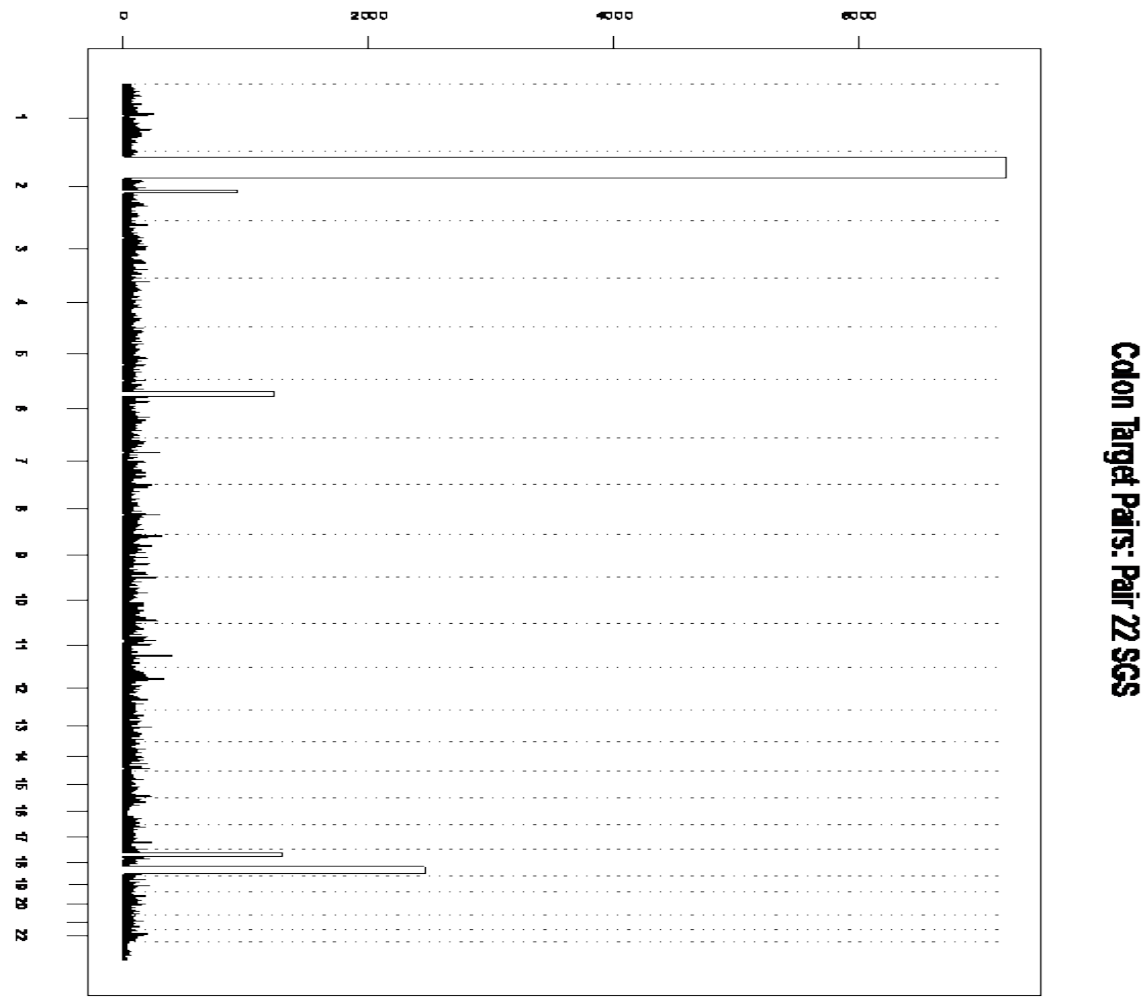


Distant relatives will share genomic regions inherited from a common ancestor, with an expected length of sharing based on distance. Sharing of longer segments than expected (and sharing a phenotype) suggests they are cosegregating.

Powerful approach for rare phenotypes

Thomas A, … Cannon-Albright LA, 2008

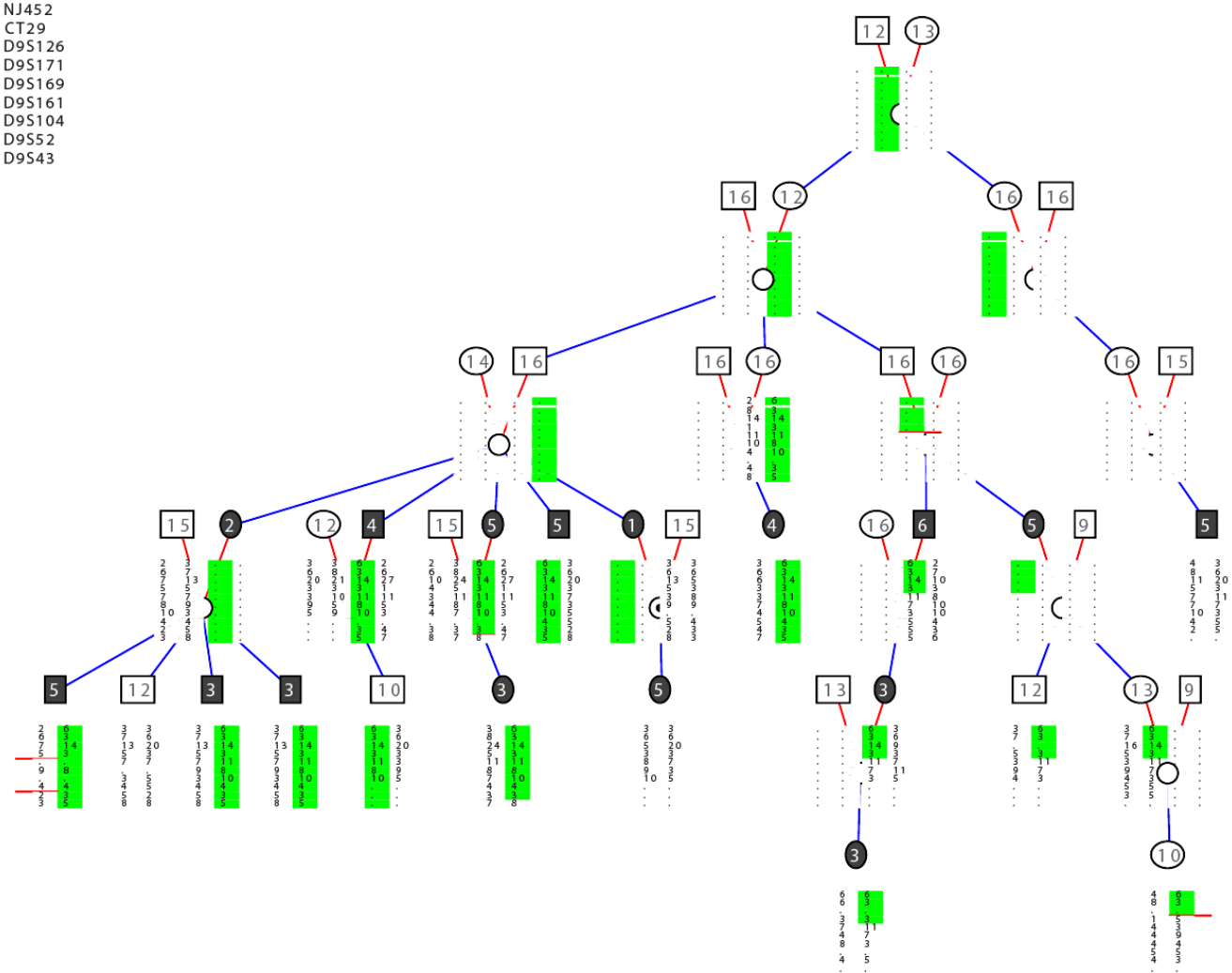
Shared Genomic Segments in Pairs of Affecteds (second cousins)



Pair shares MSH2 variant on chromosome 2

Melanoma Linkage - CDKN2A

Kindred:1771
Phenotype:
Markers:
IFN
NJ452
CT29
D9S126
D9S171
D9S169
D9S161
D9S104
D9S52
D9S43



Just sample and sequence
2 affected cousins!

Genes Localized/Isolated by Genetic Epidemiology group using UPDB high-risk pedigrees

1987	**Neurofibromatosis**
1988	**Alport Syndrome**
1994	***CDKN2A- p16*** - melanoma
1994	***BRCA1***
1996	***BRCA2***
2001	***HPC2/ELAC2*** - prostate cancer
2017	***GOLM1*** - New melanoma predisposition gene

-testing new candidates for osteoporosis, colon cancer, melanomas, diabetes now

Today's Outline

The Utah Genealogy Resource

Heritable Contribution to Parkinson's Disease

Utah Gene Discovery

Genealogies and Medicine

Genealogies and Medicine

Power and utility of genealogies for genetic studies/medicine

1. High-risk pedigree identification
   identify predisposition genes
   screen/counsel relatives for known predisposition genes

2. Risk prediction from family history or genetic data
   Apply individualized screening
   Identify high-risk populations for clinical trials

3. Gene x gene, gene x environment interaction studies

4. Inheritance, co-aggregation of multiple phenotypes in cases and their relatives

bridge from research to clinical practice

Example study of Parkinson's using UPDB

Objective: Examine association of Parkinson's Disease with cancer

Methods: use 2.3 million individuals in UPDB and linked Utah Cancer Registry

Results: Melanoma and prostate cancer were observed in significant excess among Parkinson's patients and among their relatives

Conclusions: A genetic link with Parkinson's and some cancers was strongly supported

Kareus et al Arch Neurol 2012

What are we doing today?

Sequence Utah high-risk Parkinson's pedigrees to identify predisposition genes

Use UPDB to define risk of Parkinson's based on complete family history of Parkinson's in first to third-degree relatives

SUMMARY

Unique Utah resources can result in seminal contributions to knowledge about inherited predisposition to Parkinson's Disease

Important elements are:

- UPDB genealogy resource with linked medical data
- high quality phenotypes
- sample and data collection - helpful families!!!

Questions?