Haplotyping in Families Practical

Gonçalo Abecasis

The Gene Flow Pattern

- At each marker location ℓ
- Define inheritance vector $\mathbf{v}_\ell$
  - 2^{2n} alternative outcomes
  - Meiotic outcomes specified in index bit
  - Likelihood for each gene flow pattern
    - Conditional on observed genotypes
- For haplotyping, choose most likely pattern

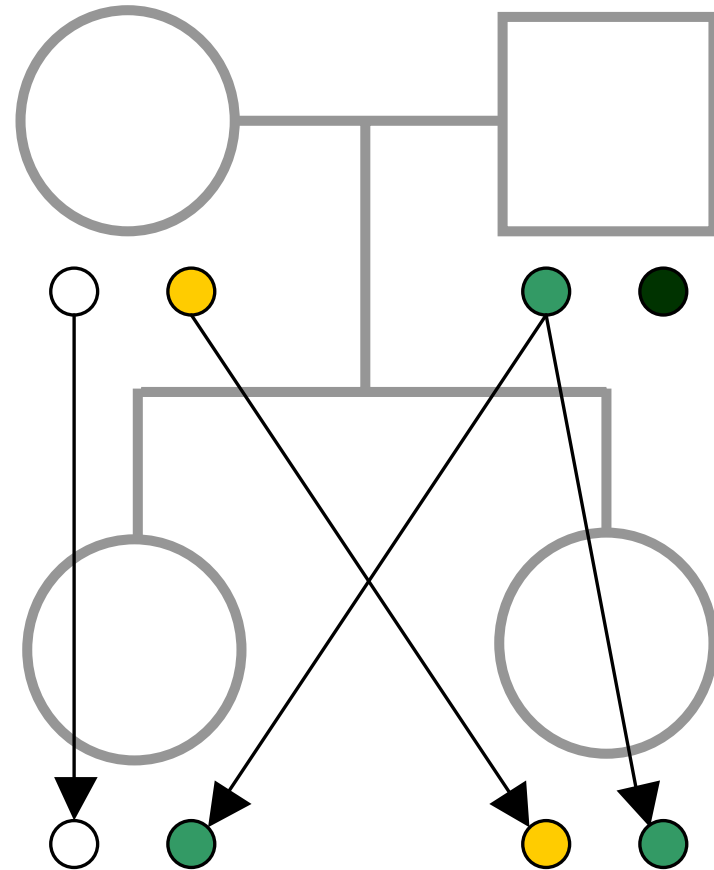| 0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L

Observed Genotypes

- For each family
- For each marker
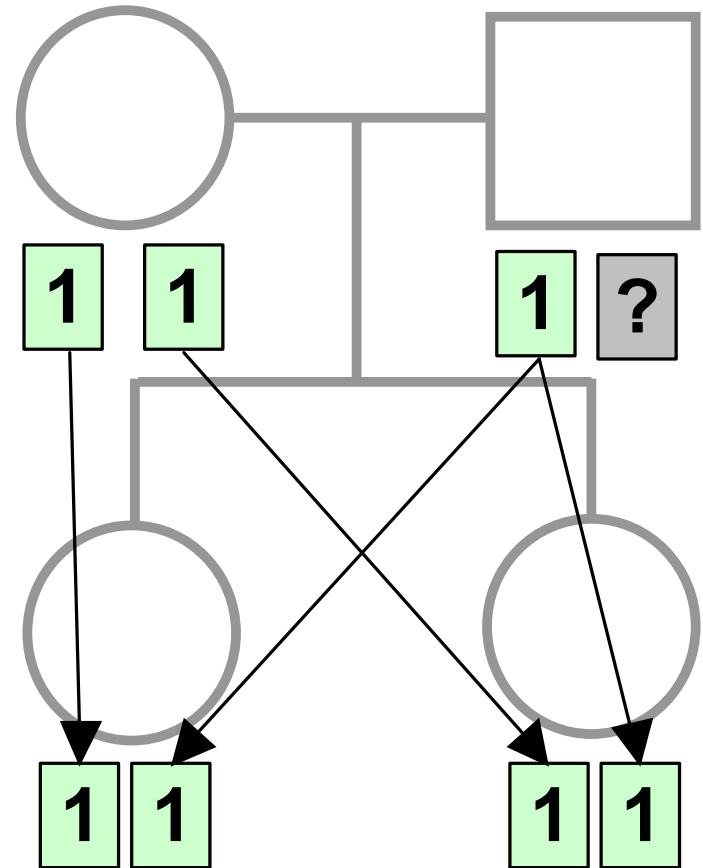
- Some pattern of observed genotypes
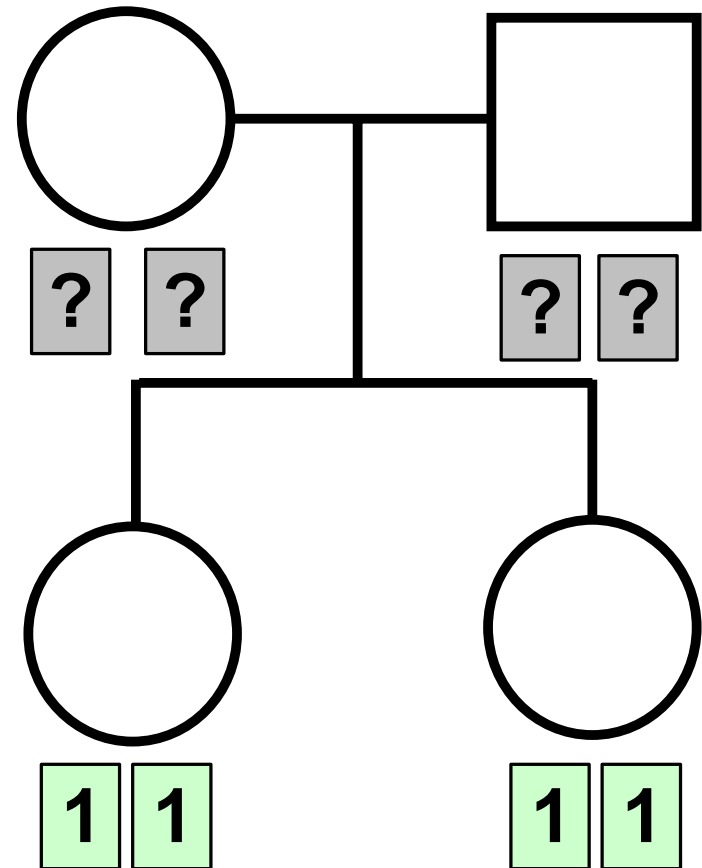
Gene flow pattern

- In turn, specify gene flow throughout the pedigree

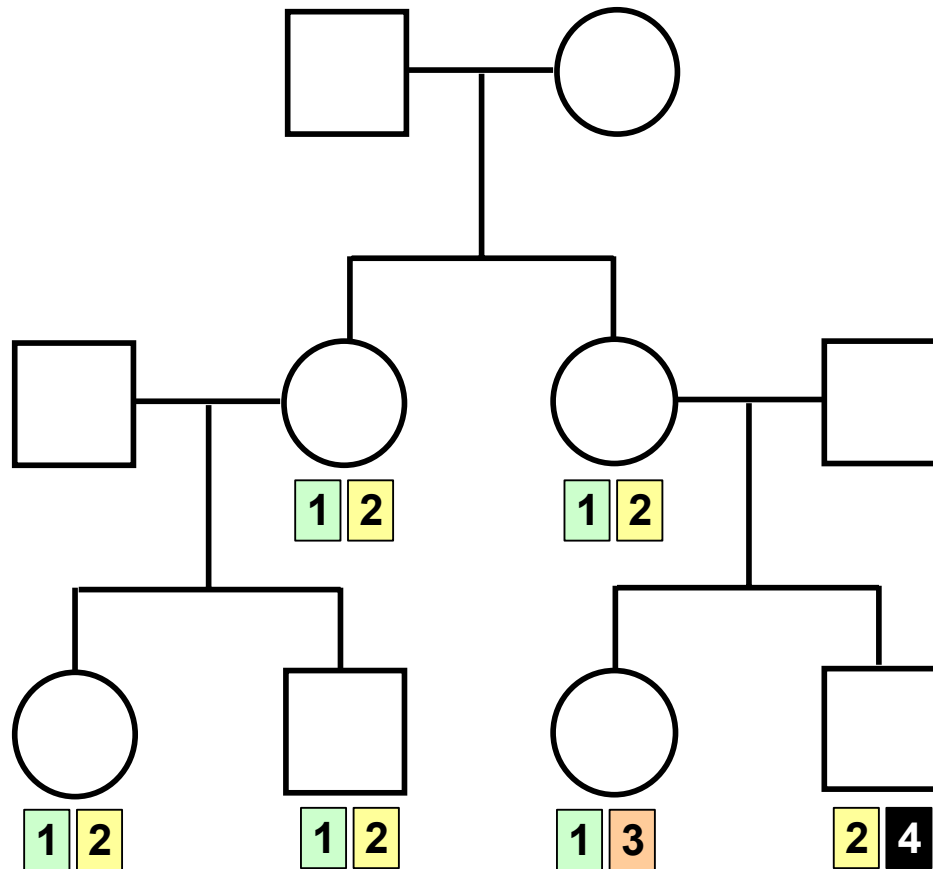- For each individual, we know precisely what founder allele they carry

Combine the two…

- Conditional on gene flow (aka, inheritance vector)

- In this case one founder allele set, {1, 1, 1, ?}

- Conditional likelihood
  - $P(\text{allele } 1)^3 \, P(\text{any allele})$

A more complex example

A more complex example …
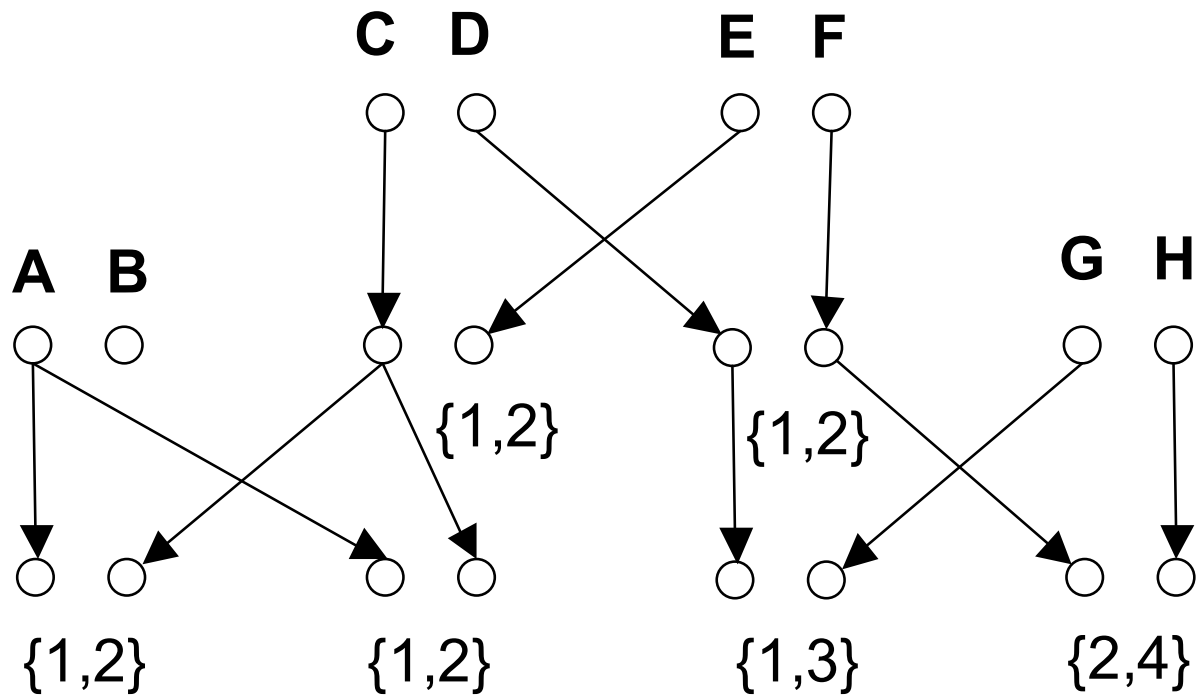
Founder Allele Graph II

- Possible groups of founder alleles
  - Any Allele Singleton
    - Passes only through untyped individuals
  - One Possible State For Each Allele
    - At least one allele passes through homozygotes or different heterozygotes
  - Two Possible States For Each Allele
    - All alleles pass through identical heterozygotes
    - Arbitrarily fixing one founder allele sets others

Possible founder allele states…

Founder Alleles in Group	Corresponding Allele States	Probability
(B)	(any allele)	1
(A,C,E)	(1,2,1) or (2,1,2)	$P(1)^2P(2)+P(2)P(1)^2$
(D,E,F,G)	(1,2,3,4)	$P(1)P(2)P(3)P(4)$
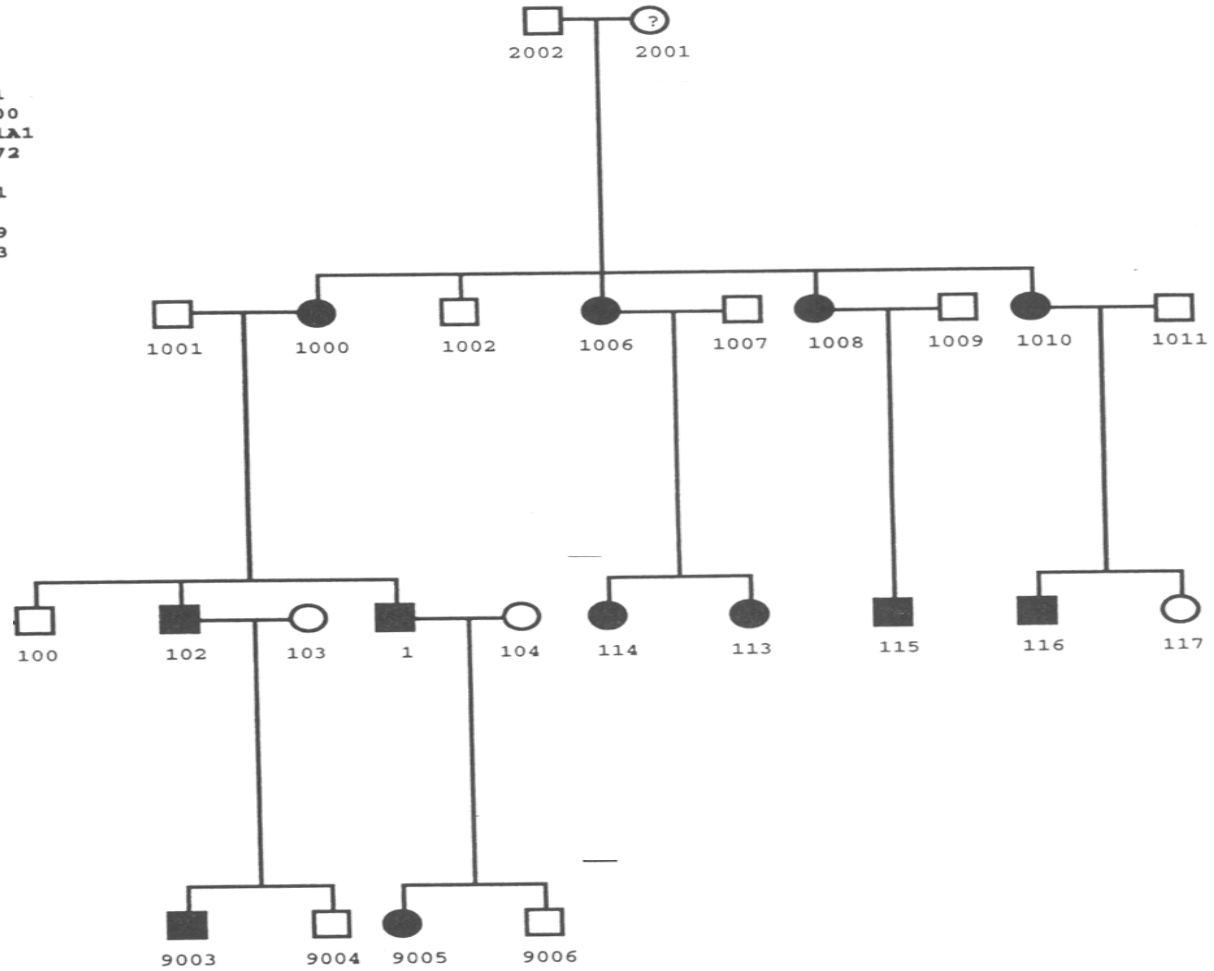
pterm.
D12S91
D12S100
CACNL1A1
D12S372
pY2/1
pY21/1
KCNA5
D12S99
D12S93
cen.

Episodic Ataxia Pedigree

Episodic Ataxia Pedigree

Marker order (pter → cen):

- pterm.
- D12S91
- D12S100
- CACNL1A1
- D12S372
- pY2/1
- pY21/1
- KCNA5
- D12S99
- D12S93
- cen.

Generation I

2002 (□) — 2001 (○ ?)

2002		2001	
1	1	2	3
4	1	5	6
10	9	7	10
2	2	3	2
1	5	1	2
7	5	10	7
3	4	4	2
6	4	5	3
7	4	4	4

Generation II

1001 (□), 1000 (●), 1002 (□), 1006 (●), 1007 (□), 1008 (●), 1009 (□), 1010 (●), 1011 (□)

1001		1000		1002		1006		1007		1008		1009		1010		1011	
1	3	3	1	2	1	3	1	3	4	3	1	2	1	3	1	3	3
3	3	6	4	5	4	6	4	6	1	6	1	6	1	6	1	4	2
9	9	10	10	7	10	10	10	8	1	10	9	3	9	10	9	8	9
5	4	3	2	2	2	3	2	3	2	3	2	3	2	3	2	3	1
6	6	2	1	1	1	2	1	2	1	2	5	1	1	2	5	6	4
3	4	7	7	10	7	7	7	7	7	7	5	5	4	7	5	5	4
3	4	2	3	4	3	2	3	4	7	2	4	1	2	2	4	4	1
1	5	3	6	5	6	3	6	2	4	3	4	2	5	3	4	4	3
6	6	7	7	4	7	7	7	3	3	4	4	2	1	4	4	4	4

Generation III

100 (□), 102 (■), 103 (○), 1 (■), 104 (○), 114 (●), 113 (●), 115 (■), 116 (■), 117 (○)

100		102		103		1		104		114		113		115		116		117	
1	1	3	3	2	1	3	1	3	4	3	3	3	4	3	1	3	3	1	3
4	3	6	3	2	3	6	3	4	9	6	6	6	4	6	3	6	2	1	4
10	9	10	9	10	10	10	9	9	9	10	10	10	1	10	9	10	9	9	8
2	5 ←	3	4	4	3	3	5	3	2	3	3	3	2 ←	3	2	3	3	2	3
1	6 ←	2	4	6	6	2	6	6	2	2	6	1	6	2	1	2	1	5	6
7	4 ←	7	4	4	3	7	3	10	5	7	4	7	4	7	4	7	4	5	5
3	3	2	4	4	1	2	3	1	1	2	2	2	2	2	2	2	1	4	4
6	1	3	5	3	4	3	1	8	7	3	3	3	4	3	5	3	3	4	3
7	6	7	6	6	4	7	6	3	6	4	4	4	4	4	1	4	4	4	4

Generation IV

9003 (■), 9004 (□), 9005 (●), 9006 (□)

9003		9004		9005		9006	
2	3	1	3	1	3	1	1
2	6	3	6	4	6	4	3
10	10	10	10	9	10	9	9
4	3	3	3 ←	3	3	3	5
6	2	6	6	6	2	6	3
4	7	3	4	5	7	5	3
3	2	1	4	1	2	1	3
4	3	6	5	7	3	7	1
6	4	4	6	6	4	6	6