

Crohn's Disease

Alan R. Rogers

March 11, 2018

Crohn's disease

- ▶ Inflammation of bowel.
- ▶ Most common in northern Europe and Ashkenazi Jews.
- ▶ 35,000 deaths in 2010.
- ▶ No cure
- ▶ Siblings of patients have 30× normal risk.
- ▶ 70 genes involved.

1 / 22

2 / 22

IBD5 haplotype

- ▶ LD across 250 kb—a region including 5 genes.
- ▶ Frequency 40% in Europe; <5% in Africa & Asia.
- ▶ Why would a deleterious allele be so common?

OCTN1 gene w/i IBD5 haplotype

- ▶ Ergothioneine: an antioxidant synthesized by fungi and present in most plants and animals.
- ▶ OCTN1 encodes a protein that transports ergothioneine.
- ▶ Allele 503F ↑ transport.
- ▶ Associated w/ Crohn's disease—but why?

3 / 22

4 / 22

Ergothioneine was rare in Neolithic diet

Table 1. Ergothioneine Content of Various Foods (Data from Ey et al. (2007)).

Food	Ergothioneine (mg/kg wet weight)
Oyster mushrooms	118.91
Garlic	3.11
Pork	1.68
Beef	1.33
Chicken	1.15
Portabella mushrooms	0.93
Wheat bran	0.84
Broccoli	0.24
Onion	0.23
Spinach	0.11
Milk	<0.01
Lentils	<0.01
Green peas	<0.01
Wheat flour (refined)	<0.01
Barley flour (refined)	<0.01

So what? Why think that ergothioneine is important?

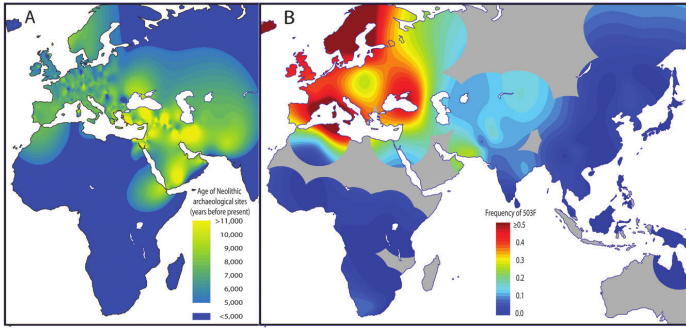
- ▶ Function poorly understood.
- ▶ Antioxidant.
- ▶ Protects against neurotoxins.
- ▶ OCTN1 is highly conserved in vertebrates but has no known function apart from transporting ergothioneine.

Hypothesis: 503F allele is a beneficial adaptation to low dietary ergothioneine but is linked to a deleterious mutation causing Crohn's disease. (Huff et al, 2012)

5 / 22

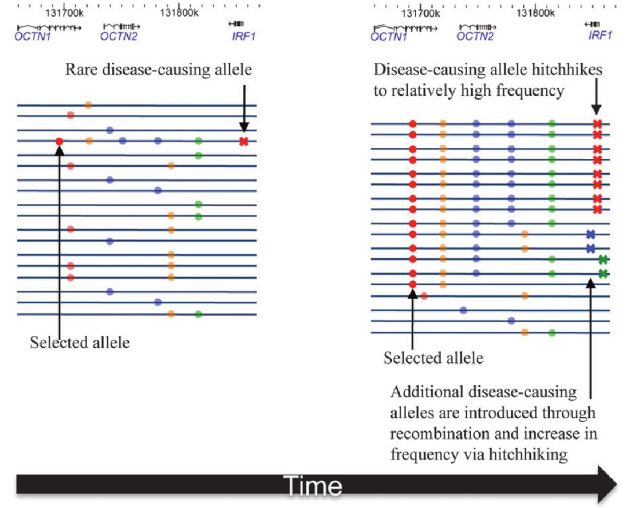
6 / 22

Map of 503F allele at OCTN1

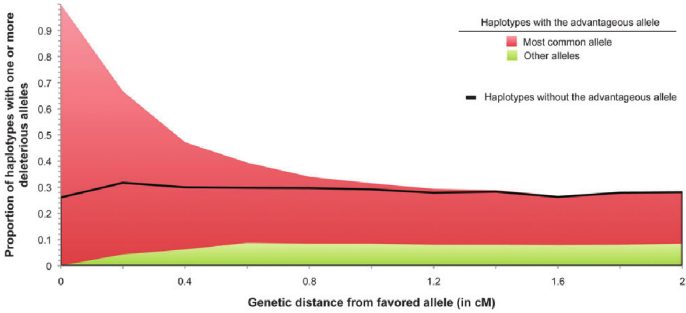


(A) age of earliest Neolithic; (B) frequency of 503F allele.

How adaptation can cause disease



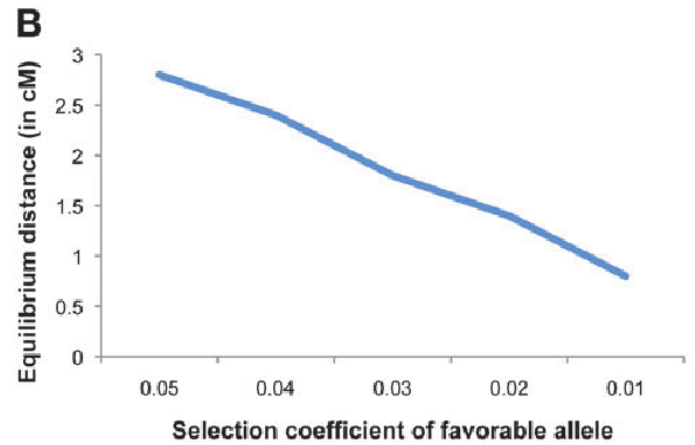
Deleterious hitchhikers are common during selective sweep



Red and green show frequency of deleterious variants among favored haplotypes.

Deleterious alleles may each be rare. Only detectable association may be that of disease with common favorable allele.

Deleterious hitchhikers may be far away if selection is strong



Common neutral mutations

- ▶ May accidentally drift to high frequency, but this takes a long time.
- ▶ Plenty of time for recombination.
- ▶ Sit on *short* stretches of original chromosome.

Common favorable mutations

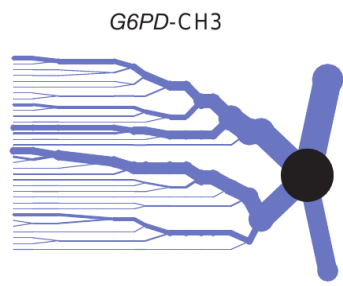
- ▶ Increase rapidly in frequency
- ▶ Little time for recombination.
- ▶ Sit on *long* stretches of original chromosome.

DNA sequences from region of human lactase gene

```

20 cgcttcaggcatttcctatctaaacagaccaacgtaAgggtacaatgcctaaccagagctttcaactct
21 .....
22 .....
23 .....
24 .....
25 .....
26 .....
27 .....t.....
28 .....t.....
29 .....G.a.g.....t.....gac.c.tgtct.
37 .....cgga...gat.at..gg..c....tc.gGaaa.g..ccttt..tg.....c...t.t...
38 .....cgga...gat.at..gg..c....tc.gGaaa.g..ccttt..tg.....c...t.t...
39 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
40 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
41 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
42 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
43 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
44 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
45 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
46 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
47 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
48 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
49 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
50 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
51 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
52 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
53 .....tcc...agtag.t.cat..g....t.ttcogG..a.gt.....t.....gac.c.tgtct.
    
```

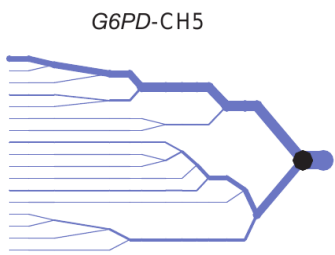
Haplotype bifurcation diagram: no selection



Horizontal axis: position on chromosome.

Simultaneous plot of LD and allele frequency
 ● "core haplotype." Diameter = frequency.
 Purple lines: haplotypes linked to core. Thickness = frequency
 Bifurcations: recombination
 Drift: lines diminish rapidly in thickness

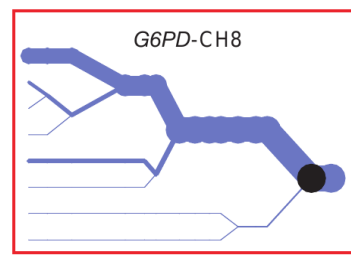
Haplotype bifurcation diagram: no selection



Small ● means core haplotype is rare.
 No evidence of selection

Horizontal axis: position on chromosome.

Haplotype bifurcation diagram: selection

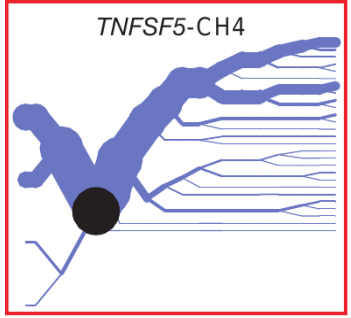
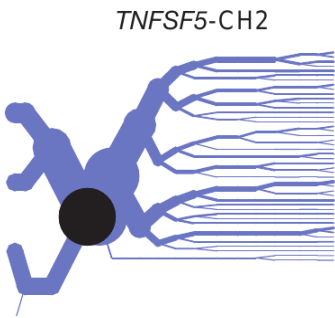


Horizontal axis: position on chromosome.

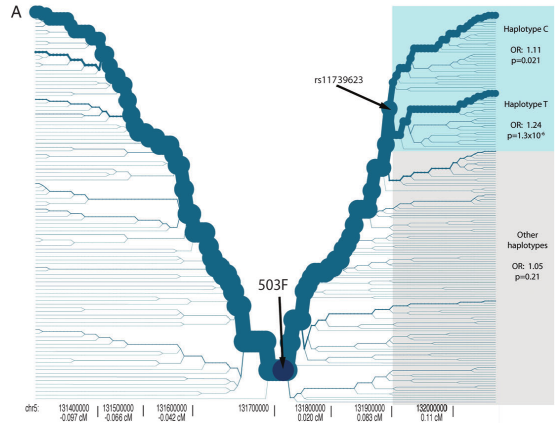
Relatively common haplotype
 Purple line stays thick: long-range LD.
 Implies selection.

No selection

Selection



Bifurcation diagram for 503F allele



Implies selection.

Date of selective sweep

Extent of LD around 503F indicates that selective sweep began 12,550 y ago (95% confidence interval: 7,750–19,025).

Wheat was domesticated 10,600 y ago, barley 9,500 y ago.

Selective sweep began during early Neolithic.

Huff et al (2012)

Testing the hitch-hiking hypothesis

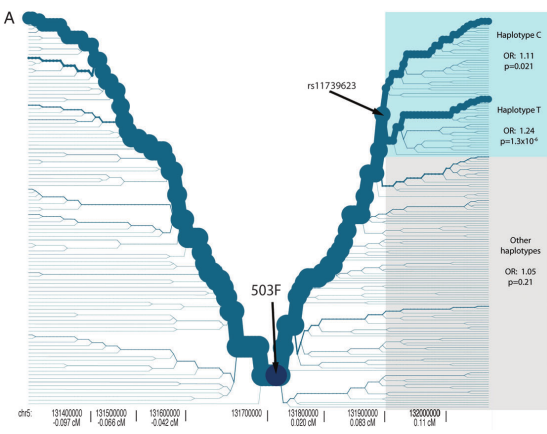
If Crohn's disease is caused by a hitch-hiker, then some 503F haplotypes may lack the disease.

Two linked genes (IRF1 and IL5) are plausible candidates for Crohn's disease.

Crohn's disease caused by hitch-hiker at IRF1 or IL5

Gray: recombination separates 503F from IRF1 and IL5. No assoc. btw 503F & Crohn's.

Blue: no recombination between 503F, IRF1, and IL5. Assoc. btw 503F & Crohn's.



Summary

Crohn's disease is the maladaptive consequence of adaptive evolution in response to Neolithic diet.