

Association whap

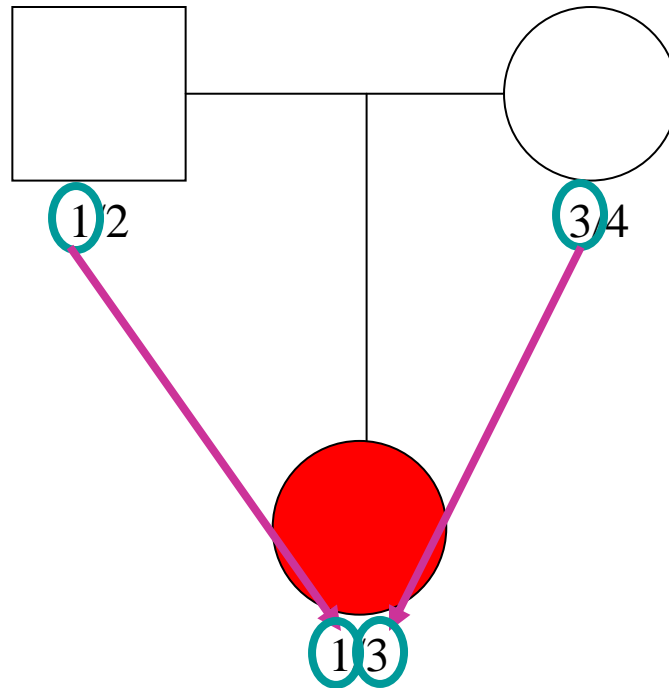
Benjamin Neale

19th International Workshop on
Twin Methodology

Practical--Whap

- Whap implements both Case/control and Family-based designs
- We will be looking at an example TDT from whap, as we have been dealing with families all week long
- Whap is found at <http://pngu.mgh.harvard.edu/~purcell/whap/>
- Copy all files from ben/association/practical

Reminder of TDT



Only transmissions from heterozygous parents at the allele of interest provide information

Under the null, we expect equal chance of either allele transmitting from a heterozygote parent to the offspring

Though Whap only does biallelic markers (SNPs)

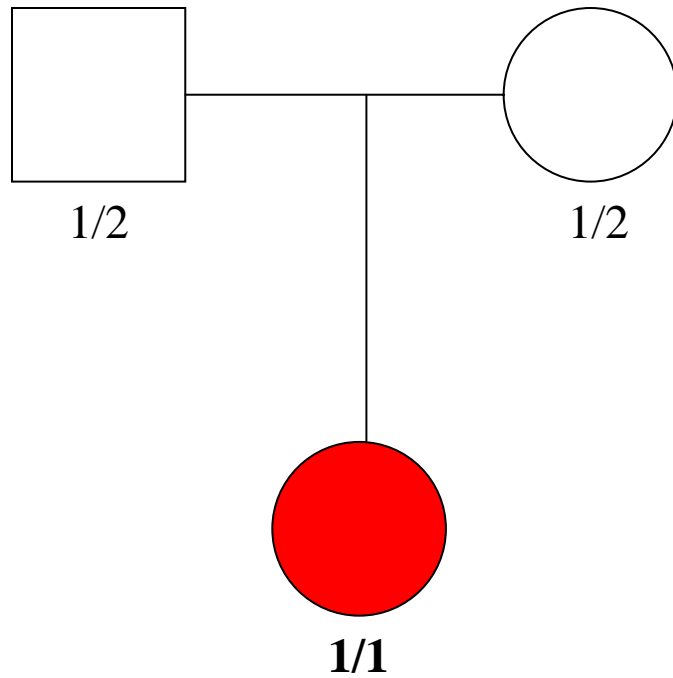
Testing in Whap

- Whap demands you specify the alternative hypothesis
- The null hypothesis can be supplied, but the default is that no marker has any effect on the trait
- When you specify an alternative hypothesis you are testing that marker/haplotype is associated with disease

Significance in Whap

- Whap uses permutation analysis to assess significance
- Permutation analysis assumes that all data come from the null distribution
- Thus, under TDT, the null distribution is that it is equally likely allele 1 or allele 2 is transmitted.
- We 'mix up' the transmitted and non-transmitted allele within family

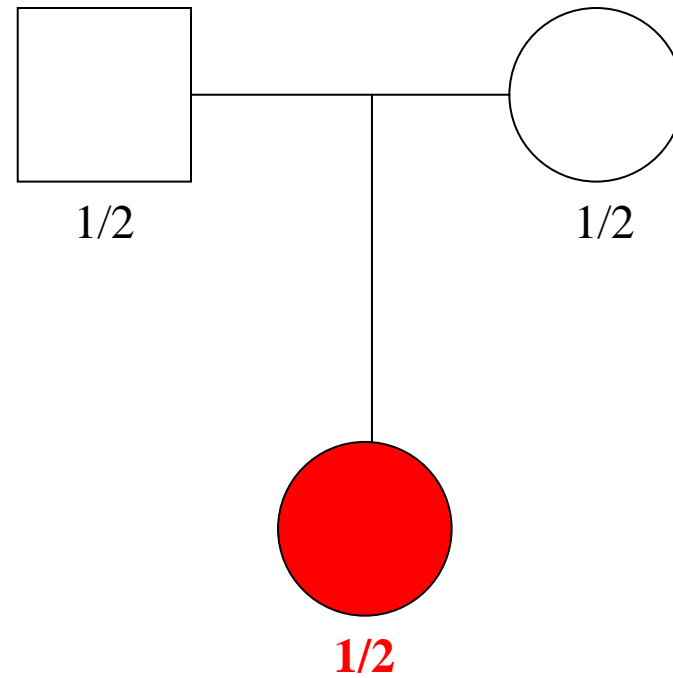
Visualization of permutation



Original data:

2: allele 1 trans

0: allele 2 trans

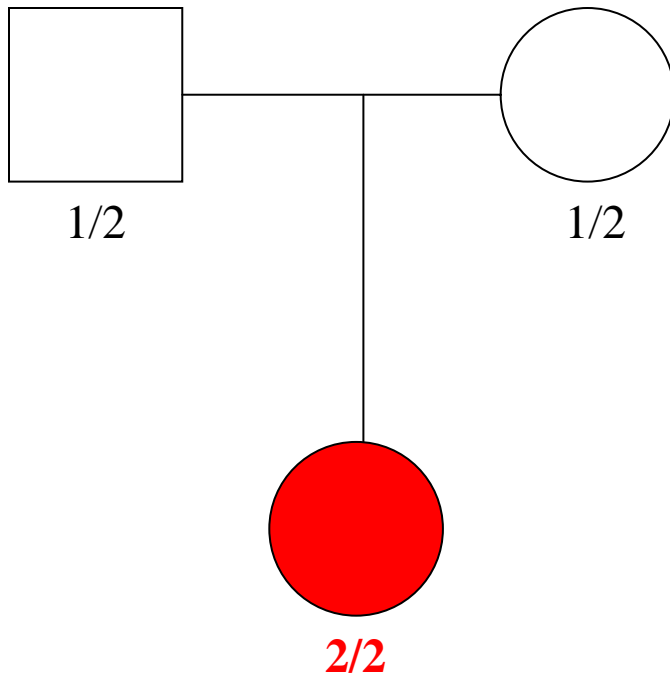


Permutation 1:

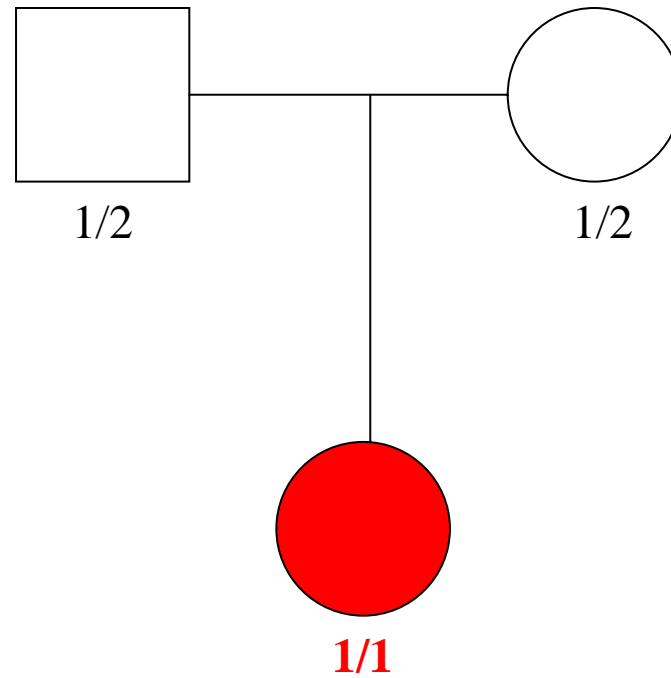
1: allele 1 trans

1: allele 2 trans

Visualization of permutation



Permutation 2:
0: allele 1 trans
2: allele 2 trans



Permutation 3:
2: allele 1 trans
0: allele 2 trans

Whap input

- 3 files
 - Ped file
 - Map file
 - Dat file

Ped files

- Pedigree file looks like:

| Family | ID | Pa | Ma | sex | Aff | SNP1 | |
|--------|-----|-----|-----|-----|-----|------|-----|
| | | | | | | AI1 | AI2 |
| PED054 | 430 | 0 | 0 | 1 | 1 | 1 | 2 |
| PED054 | 412 | 430 | 431 | 2 | 2 | 1 | 2 |
| PED054 | 431 | 0 | 0 | 2 | 1 | 2 | 2 |

The DATA file

Must have .ped extension

Founders

Map Files

| Chr | Marker name | cM | bp |
|-----|-------------|--------|--------|
| 5 | IGR1118a_1 | 100.00 | 274044 |
| 5 | IGR1119a_1 | 100.00 | 274541 |
| 5 | IGR1143a_1 | 100.00 | 286593 |
| 5 | IGR1144a_1 | 100.00 | 287261 |
| 5 | IGR1169a_2 | 100.00 | 299755 |

Tells Whap where the markers are on the genome
Must have .map extension

Dat file

| Type | Name |
|------|------------|
| A | disease |
| M | IGR1118a_1 |
| M | IGR1119a_1 |
| M | IGR1143a_1 |
| M | IGR1144a_1 |

A=affection

B= disease

T=trait

C=covariate

M=marker

S=marker skip

X=trait skip

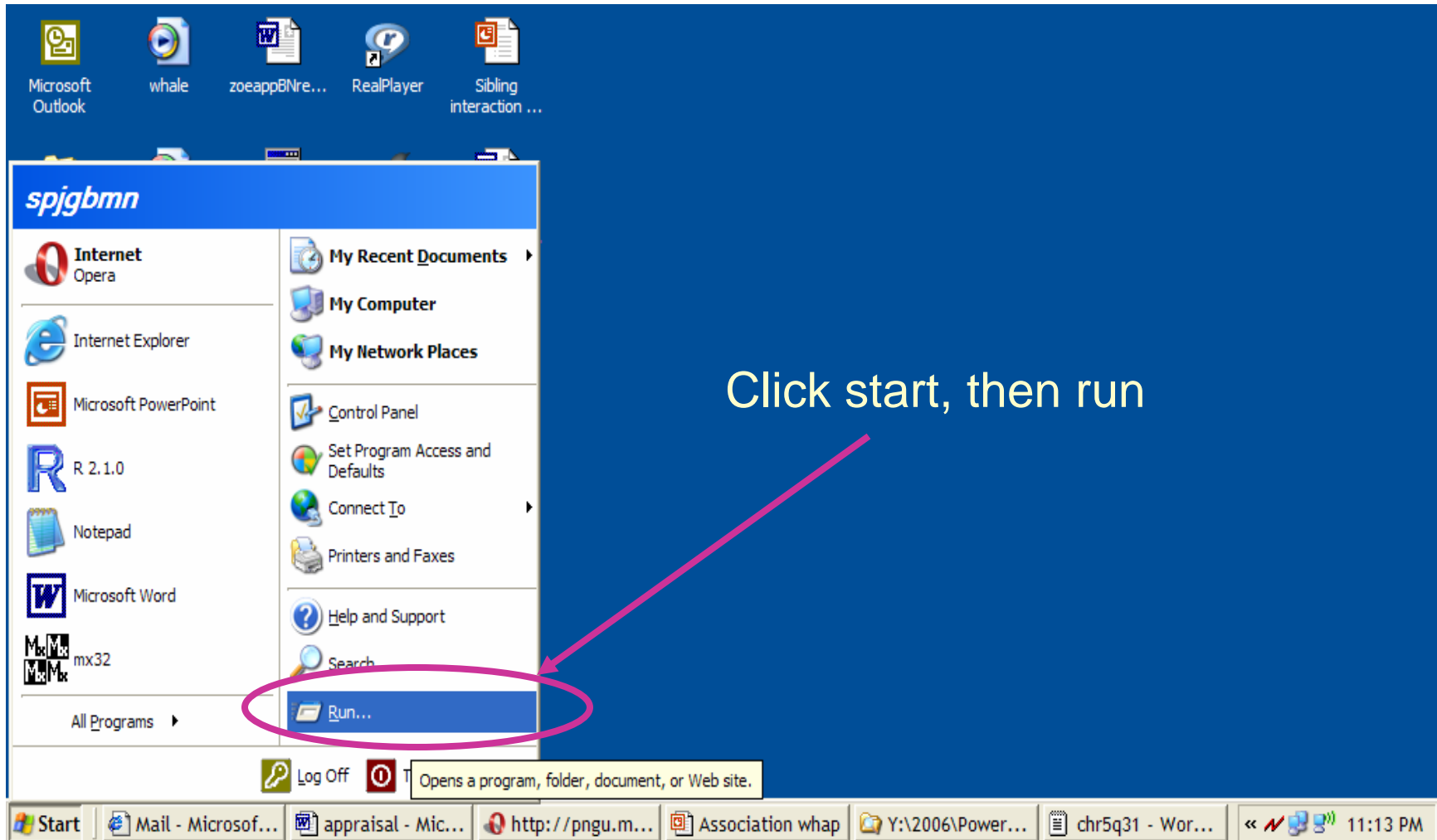
Tells Whap what is in each column

Must have .dat extension

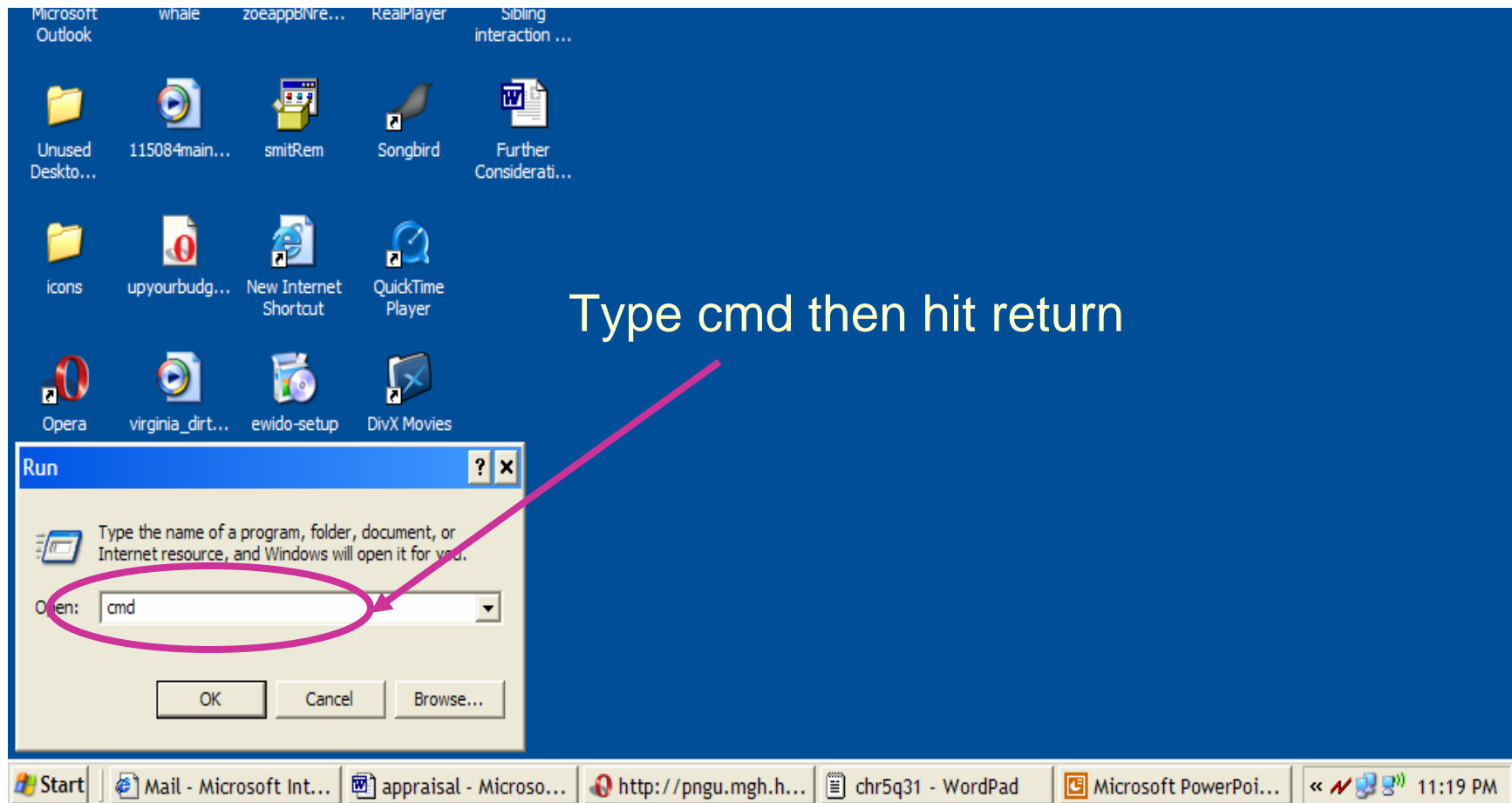
Running Whap!

- Whap is a command line program which means we have to use the command prompt in Windows.
- To run command prompt click start, then run...
- In the run box type cmd, then hit return
- Change to your network drive by typing 'H:' (not necessary if you are in H:)
- Change directory with cd 'directory name'
- Dir brings up a list of files and directories.

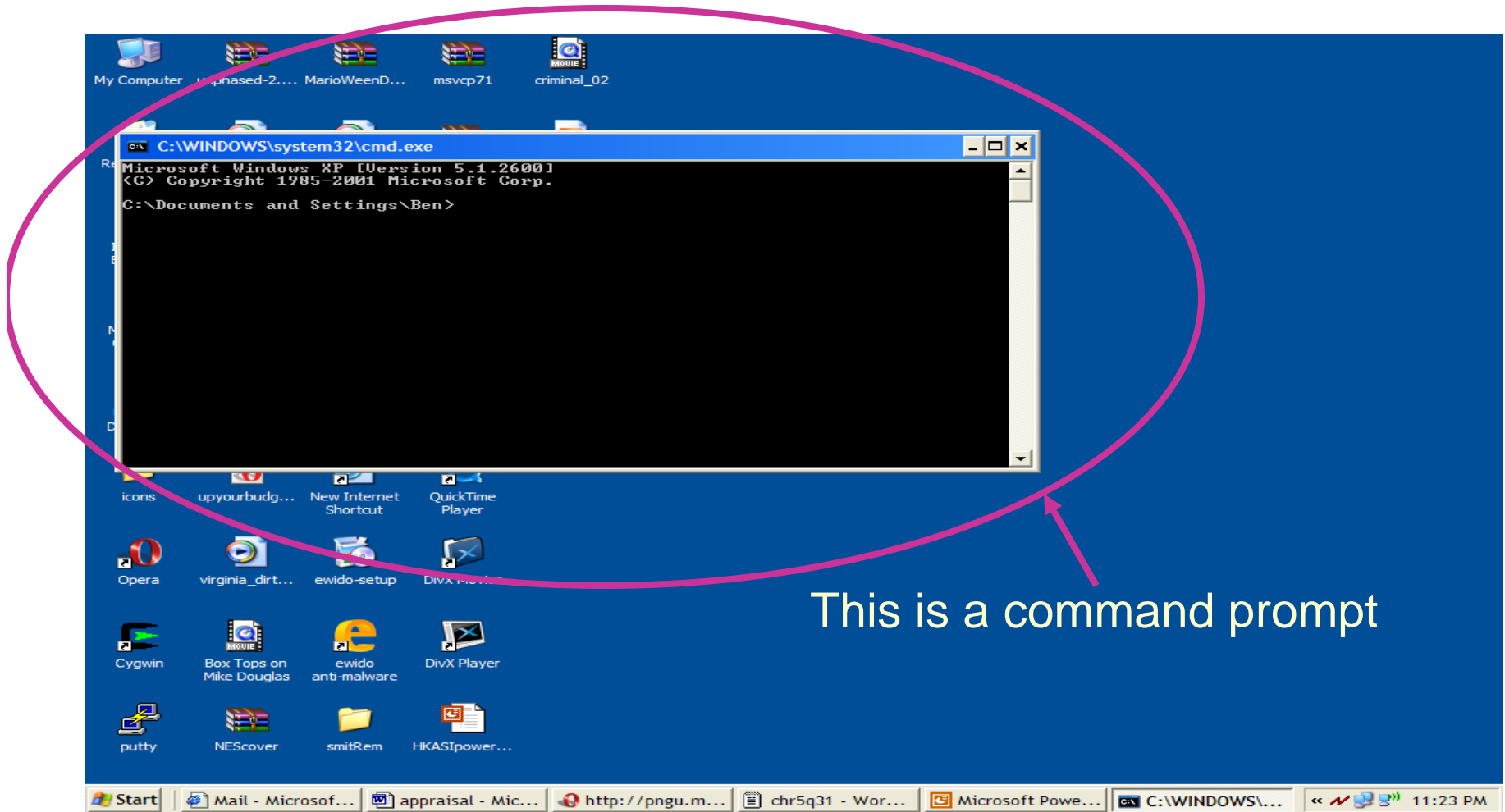
Command prompt



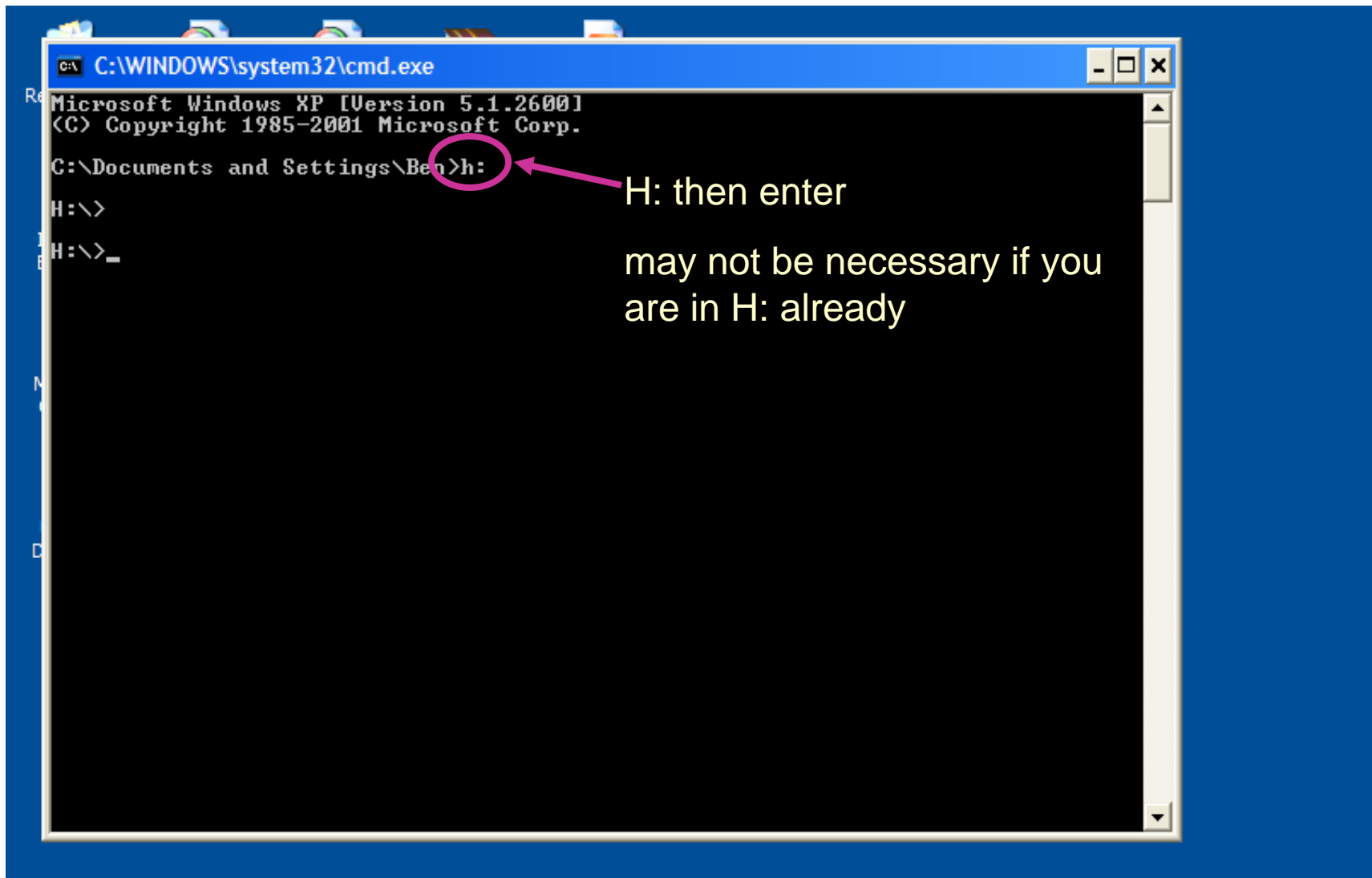
Running command prompt



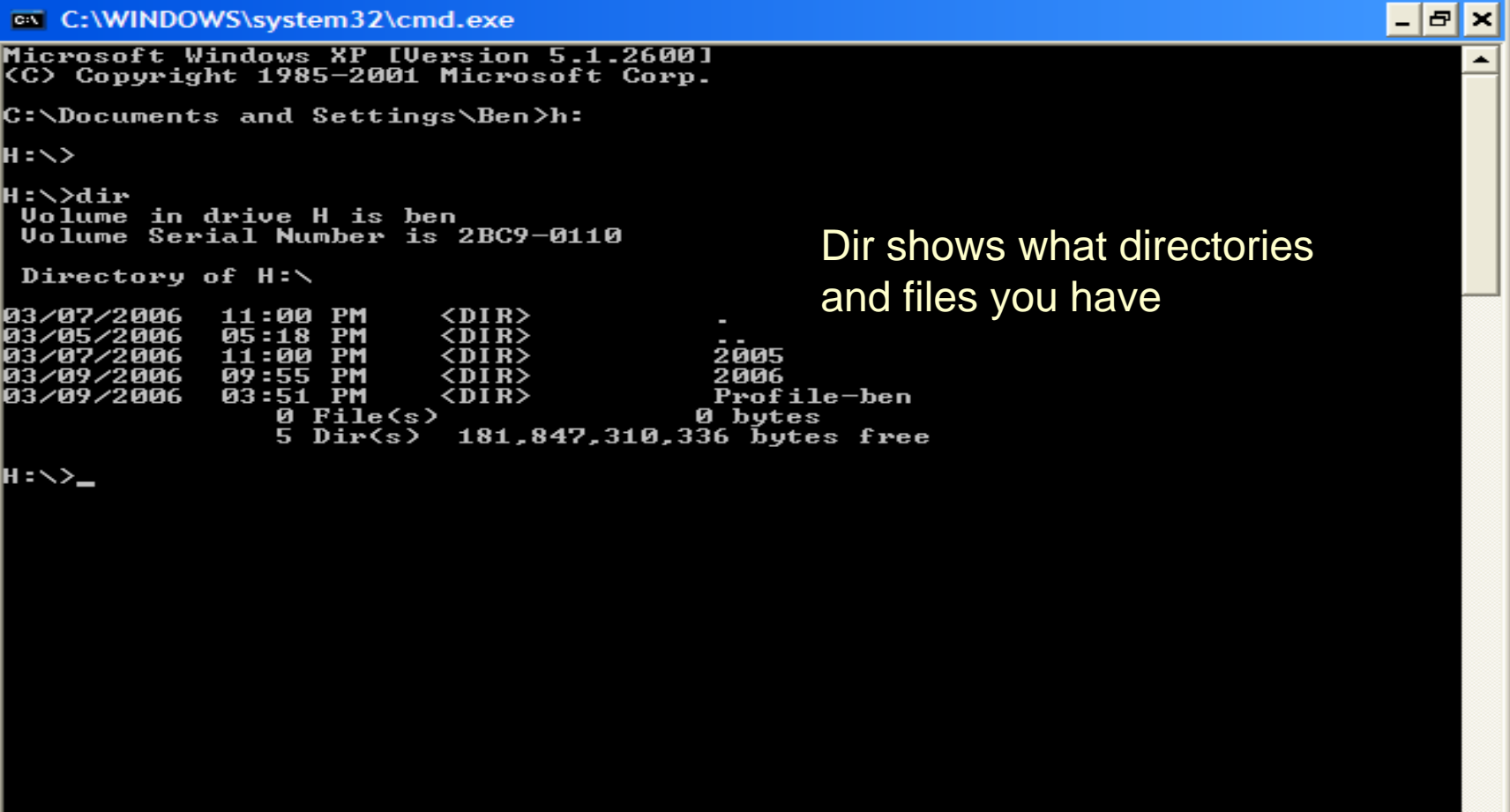
You should have this up now



Changing to H:



Dir shows you what is in your directory



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\Ben>h:
H:\>
H:\>dir
Volume in drive H is ben
Volume Serial Number is 2BC9-0110

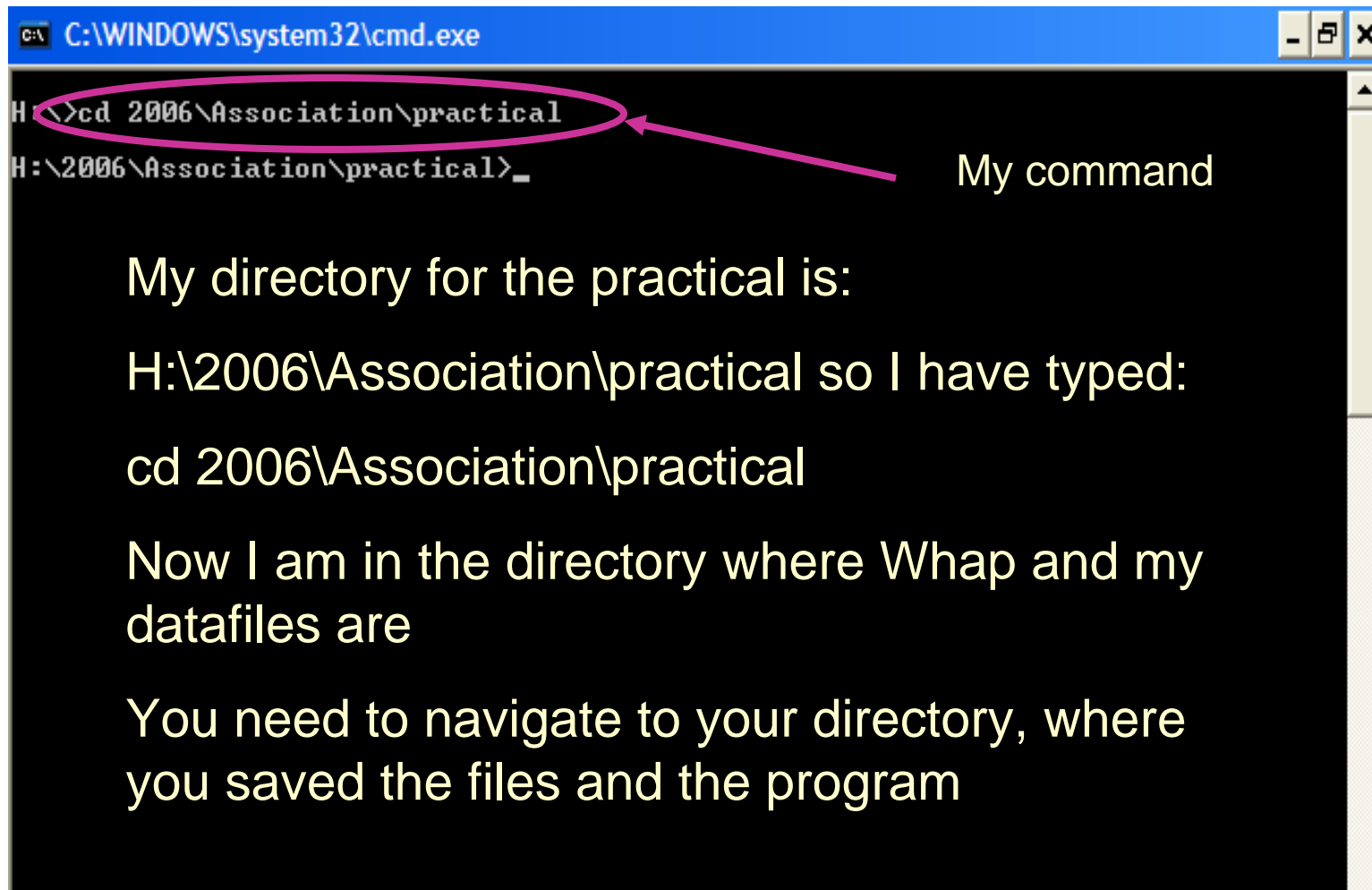
Directory of H:\

03/07/2006  11:00 PM    <DIR>          .
03/05/2006  05:18 PM    <DIR>          ..
03/07/2006  11:00 PM    <DIR>          2005
03/09/2006  09:55 PM    <DIR>          2006
03/09/2006  03:51 PM    <DIR>          Profile-ben
           0 File(s)                0 bytes
           5 Dir(s)  181,847,310,336 bytes free

H:\>_
```

Dir shows what directories and files you have

My directory is...



The image shows a screenshot of a Windows command prompt window. The title bar at the top reads "C:\WINDOWS\system32\cmd.exe". The command prompt shows the current directory as "H:\>" and the command ">cd 2006\Association\practical" has been entered. A pink oval highlights the command, and a pink arrow points from the text "My command" to the command. Below the command prompt, the text "My directory for the practical is:" is followed by "H:\2006\Association\practical so I have typed:" and the command "cd 2006\Association\practical". The text "Now I am in the directory where Whap and my datafiles are" and "You need to navigate to your directory, where you saved the files and the program" follows.

```
C:\WINDOWS\system32\cmd.exe
H:\>cd 2006\Association\practical
H:\2006\Association\practical>_
```

My command

My directory for the practical is:
H:\2006\Association\practical so I have typed:
cd 2006\Association\practical

Now I am in the directory where Whap and my datafiles are

You need to navigate to your directory, where you saved the files and the program

Now what?

- Hopefully you have successfully navigated into the directory (you can check with `dir!`)
- We can now start to run Whap:
- Command is:
- `whap --file chr5q31 --alt 1 --window --cond --prev 0.01 --model w --wperm 500`

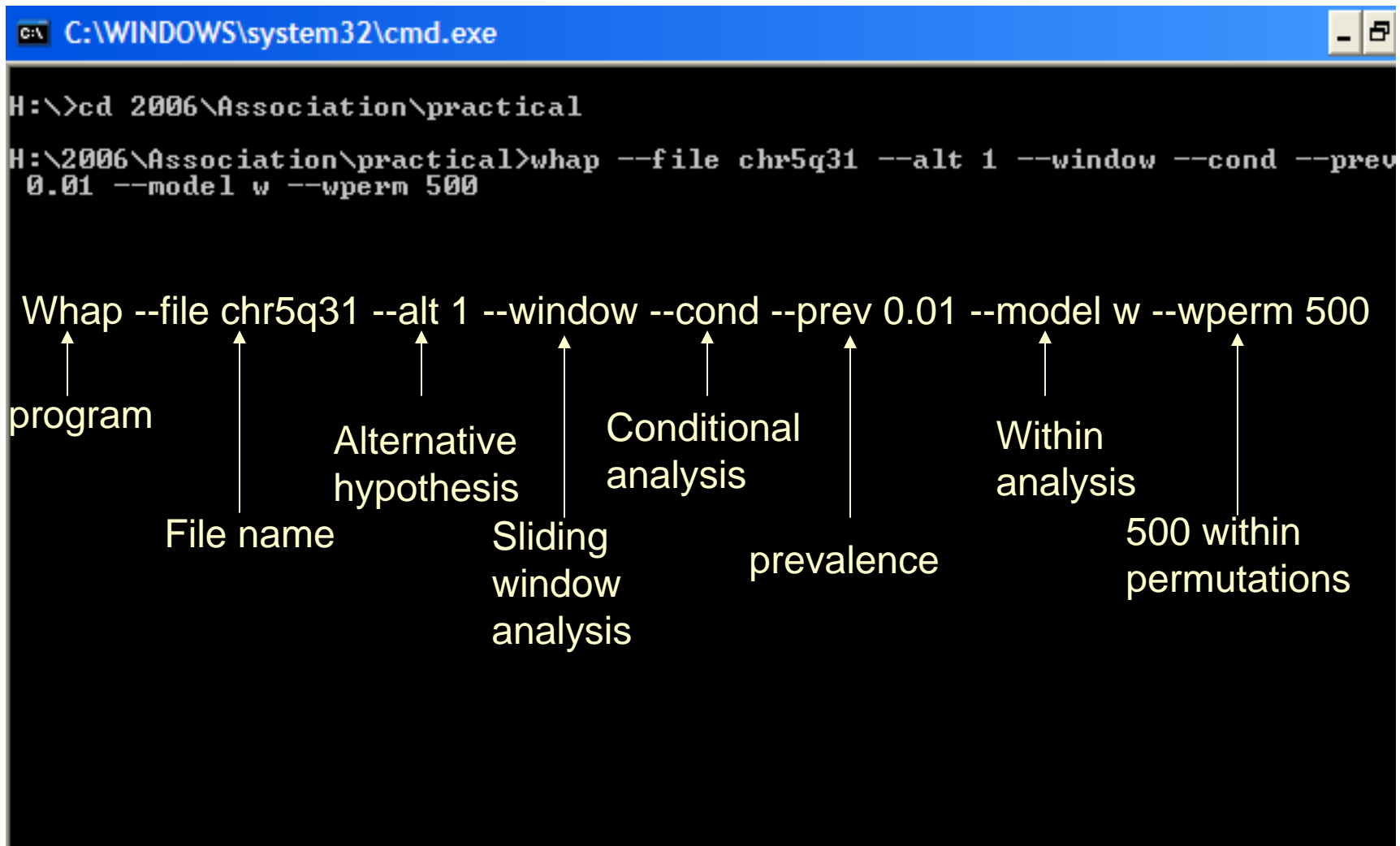
What are all these bits?

- Whap: program name
- --file chr5q3: whap requires your map, dat, and ped files to have the same name. File supplies this
- --alt 1: this is the alternative hypothesis
- --window means that whap will look at a sliding window of SNPs
 - Whap begins with the alt, then slides across all remaining markers if window is active

What are all these ‘and pieces’?

- --cond: conditional analysis—necessary for family data
- --prev 0.01: prevalence of the trait
- --model w: within model of association
- --wperm 500: number of within permutations

A picture of the command to be run



```
C:\WINDOWS\system32\cmd.exe

H:\>cd 2006\Association\practical
H:\2006\Association\practical>whap --file chr5q31 --alt 1 --window --cond --prev
0.01 --model w --wperm 500
```

Whap --file chr5q31 --alt 1 --window --cond --prev 0.01 --model w --wperm 500

↑ program

↑ File name

↑ Alternative hypothesis

↑ Sliding window analysis

↑ Conditional analysis

↑ prevalence

↑ Within analysis

↑ 500 within permutations

Output

```
C:\WINDOWS\system32\cmd.exe
WHAP! | v2.06 | 13/Dec/04 | S. Purcell, P. Sham | spurcell@pngu.mgh.harvard.edu
0 individuals w/out parents. 129 individuals with parents. Binary trait: Prevalence fixed to 0.010.
Window position : [1]
Window position : [2]
129 of 129 individuals/trios are informative
Window position : [3]
129 of 129 individuals/trios are informative
129 of 129 individuals/trios are informative
Results of sliding window analysis
Global permutation tests
-----
P_MAX = 10.217 p = 0.0059880240
P_SUM = 17.307 p = 0.0059880240
Local permutation tests
-----
>> IGR3005a_2 571022 P_1= 10.217 p= 0.0039920160
>> IGR3016a_1 576586 P_2= 7.042 p= 0.0099800399
>> IGR3018a_2 577141 P_3= 2.049 p= 0.0978043912
C:\Boulder\2006\t>
```

P max is a comparison of the best statistic you observed in your data against the 'best' statistic in each permutation

Output

```
C:\WINDOWS\system32\cmd.exe
WHAP! | v2.06 | 13/Dec/04 | S. Purcell, P. Sham | spurcell@pngu.mgh.harvard.edu
0 individuals w/out parents. 129 individuals with parents. Binary trait: Prevalence fixed to 0.010.
Window position : [1]
Window position : [2]
129 of 129 individuals/trios are informative
Window position : [3]
129 of 129 individuals/trios are informative
129 of 129 individuals/trios are informative
Results of sliding window analysis

Global permutation tests
-----
P_MAX = 10.217    p = 0.0059880240
P_SUM = 19.307    p = 0.0059880240

Local permutation tests
-----
>> IGR3005a_2 571022 P_1= 10.217    p= 0.0039920160
>> IGR3016a_1 576586 P_2= 7.042    p= 0.0099800399
>> IGR3018a_2 577141 P_3= 2.049    p= 0.0978043912

C:\Boulder\2006\t>
```

P sum is a comparison of the sum total of the statistics you observed in your data against the sum total in each permutation

Final notes

- P-sum and P-max are corrected for multiple testing, as long as you include all markers tested
- We can use Whap to do haplotype tests as well as single markers
- Much more on association at the advanced workshop