

**Summary data on fQTLs calculated from 12 cell types [using TPM]  
 These 12 cell types have at least 30% of non-zero values across all individuals**

<u>#Cell_type</u>	<u>num_fQTLs</u>	<u>min_bonf_corrected_p_val</u>	<u>max_bonf_corrected_p_val</u>
Microglia	2528	1.77E-18	0.049922298
Ex8	995	1.06E-35	0.049858929
Ex3	324	5.67E-20	0.048114878
Ex4	195	2.06E-16	0.04814516
Astrocytes	58	3.76E-06	0.045980094
Ex5	3	2.79E-06	0.00802804
In6	21	9.56E-09	0.042765232
In8	20	0.000143941	0.021903179
Endothelial	42	0.003364832	0.047003944
Ex1	0	na	na
Oligodendrocytes	0	na	na
Neuron	0	na	na

## Select ideal set of cell types

Notes:

- + Not ideal that we have 0s at bottom (and others w/very high numbers).
- + Try to stick to small set of cell types (~12ish)

Start w/original file from XS (cellfractions\_tissue\_24\_recalTPM.txt)

Exclude the cell type "Neuron"

Extract only those samples for which genotype data is available

Remove 'sparse' cell types: cases w/0 fractions in many samples

Each cell type should have appreciable diversity across individuals

--> if 0 fQTLs for a certain cell type - then that cell type does not exhibit appreciable diversity across individuals!

Real cell types rather than 'mixtures-of-cells' types.

## Remove 'sparse' cell types: cases w/0 cell fractions in many samples

cell_type	fract_non_zero_vals	mean_val	mean_val_among_non_zeros
Endothelial	0.999	0.08325	0.08331
Astrocytes	0.993	0.21417	0.21565
Ex1	0.985	0.14207	0.14425
Oligodendrocytes	0.964	0.16825	0.1745
Ex4	0.919	0.04709	0.05122
Ex5	0.886	0.05881	0.06635
Microglia	0.862	0.04531	0.05259
Ex3	0.701	0.03063	0.04368
In8	0.448	0.01	0.02229
Ex8	0.363	0.00868	0.02391
In6	0.333	0.00411	0.01233
Ex7	0.063	0.00094	0.01495
Fetal-replicating	<b>0.052</b>	0.00881	<b>0.1705</b>
Fetal-quiescent	<b>0.048</b>	0.01504	<b>0.31651</b>
OPC	<b>0.048</b>	0.00051	<b>0.01077</b>
Ex6	0.03	0.00181	0.05968
In3	0.019	0.00027	0.01442
In4	0.019	0.00031	0.01667
In5	0.019	0.00044	0.02268
In1	0.008	8.00E-05	0.01006
Ex2	0.006	5.00E-05	0.00981
In2	0	0	na
In7	0	0	na

## Select many subsets of cell types (of different sizes)

$\binom{12}{12} = 1$	12 -- just take all 12 cell types
$\binom{12}{11} = 12$	11 -- select all 12 different size-11 subsets of cell types (out of the original 12)
$\binom{12}{10} = 66$	10 -- select 10 size-10 subsets of cell types (out of the original 12)
$\binom{12}{9} = 220$	9 -- select 10 size-9 subsets of cell types (out of the original 12)
$\binom{12}{8} = 495$	8 -- select 10 size-8 subsets of cell types (out of the original 12)
$\binom{12}{7} = 792$	7 -- select 8 size-7 subsets of cell types (out of the original 12)

*Note – the subsets are not totally random. I deliberately tried to avoid cell types w/low fQTL counts (from previous analyses)*

*Ex1*

*Oligodendrocytes*

*Ex5*

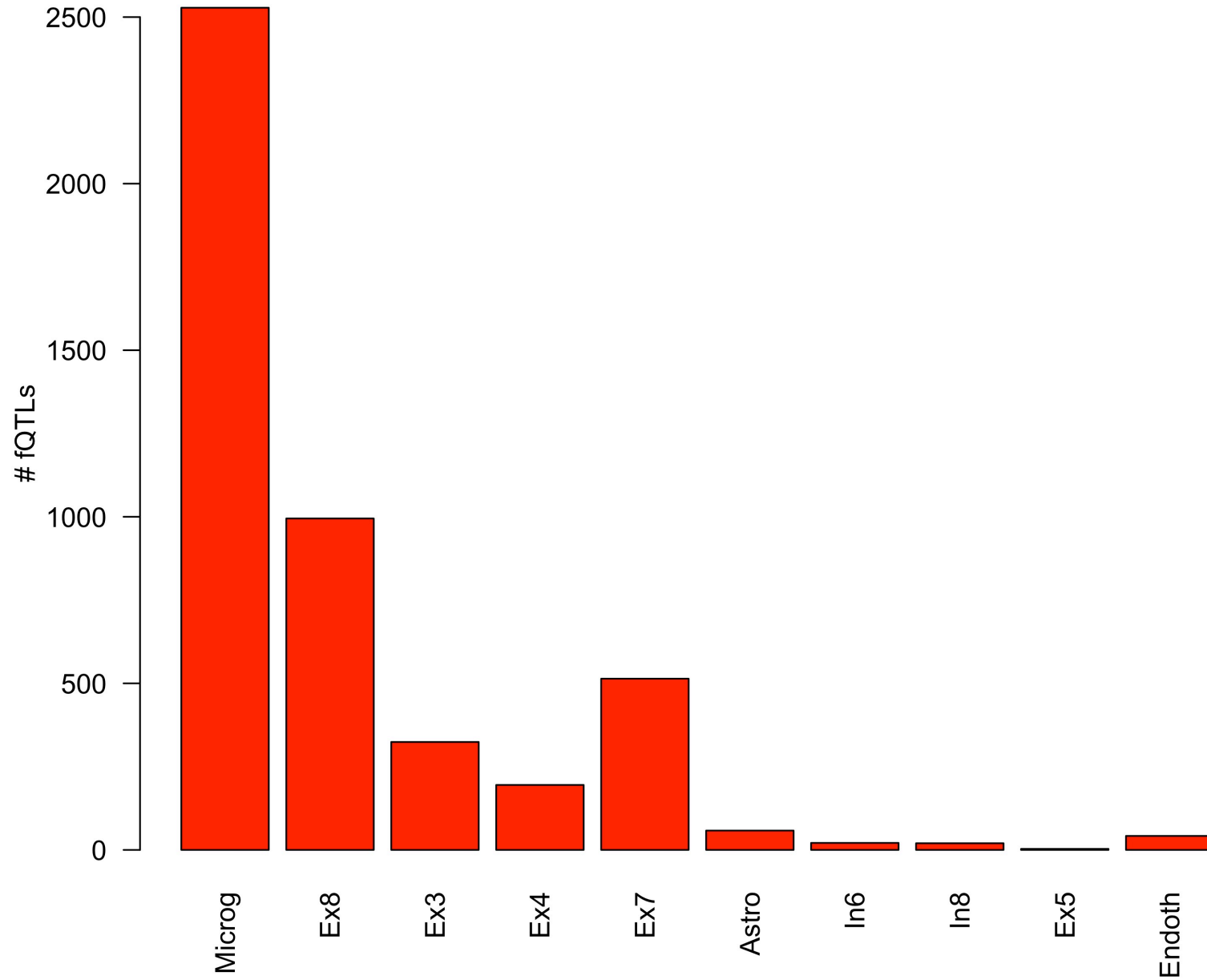
*In8*

*In6*

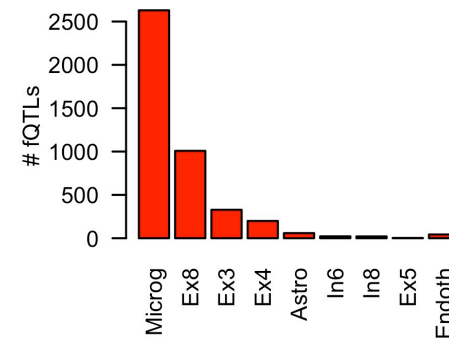
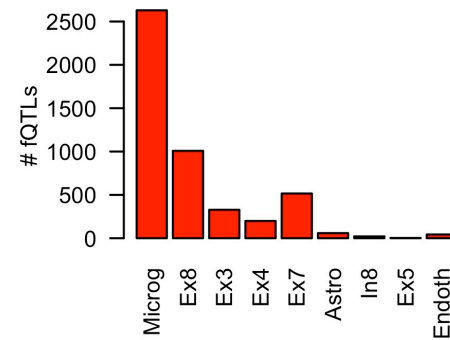
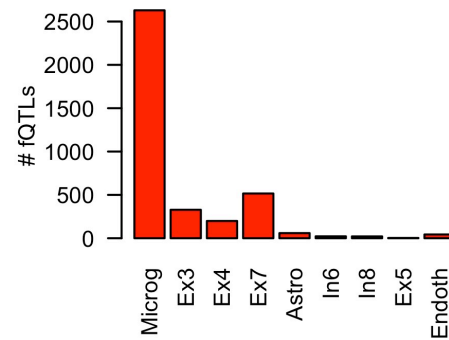
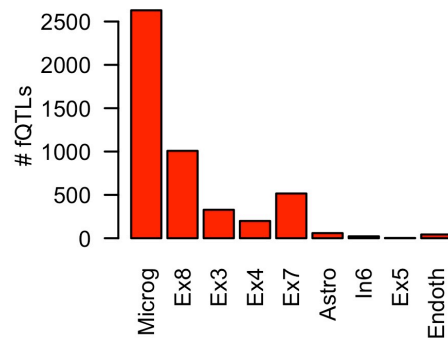
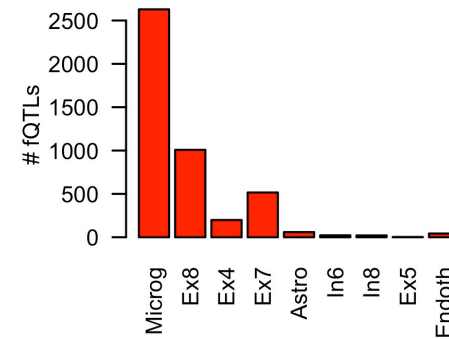
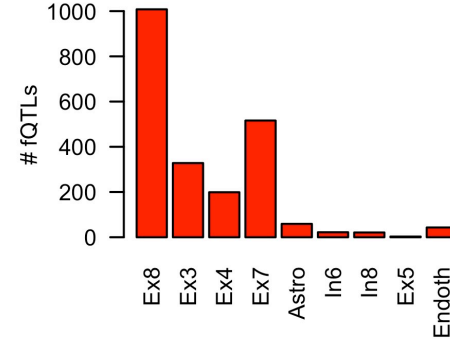
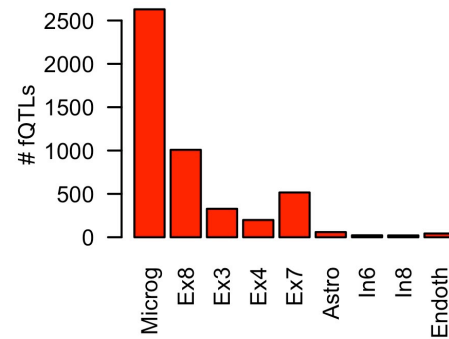
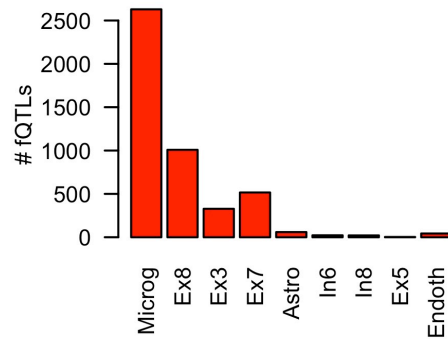
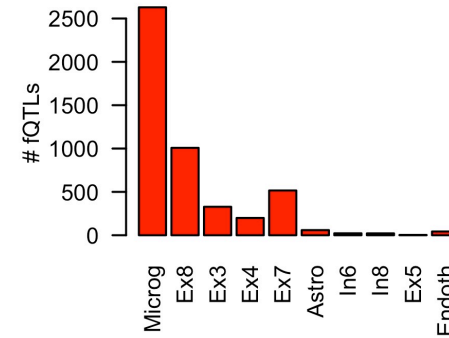
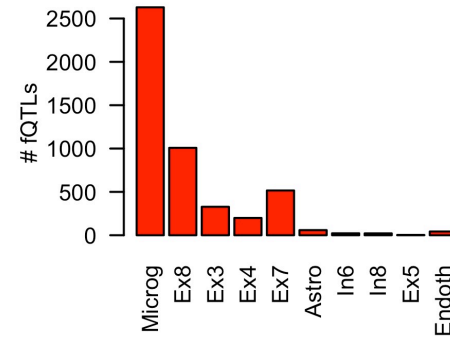
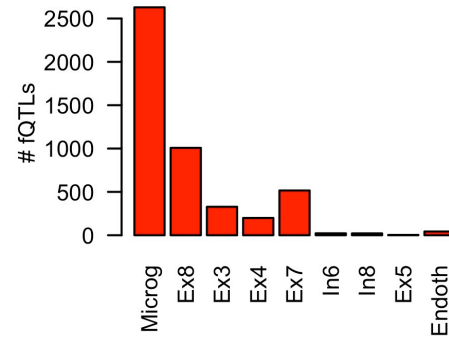
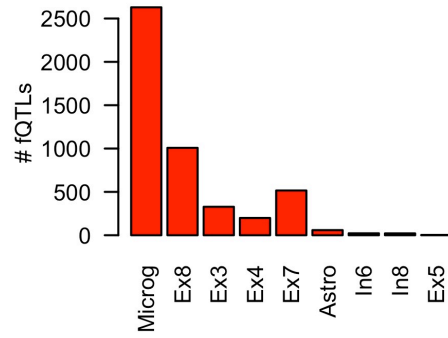
For each of the given set of call type subsets - identify fQTLs

#_phenotypes	#_SNVs	#_tests	bonf_p_val_for_0p05
12	5312628	63751536	7.84E-10
11	5312628	58438908	8.56E-10
10	5312628	53126280	9.41E-10
9	5312628	47813652	1.05E-09
8	5312628	42501024	1.18E-09
7	5312628	37188396	1.34E-09

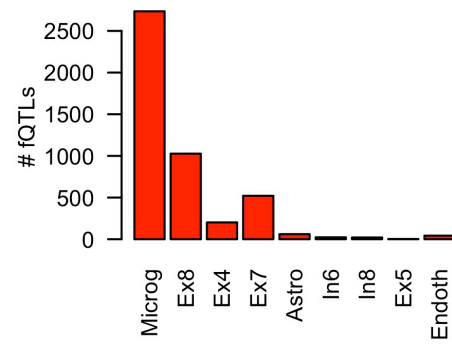
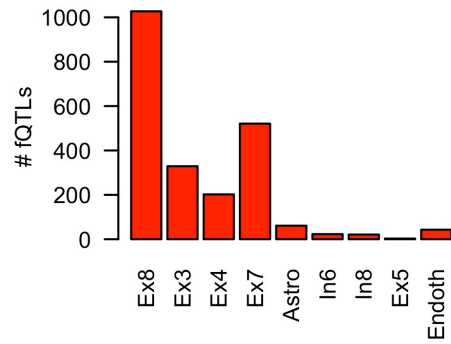
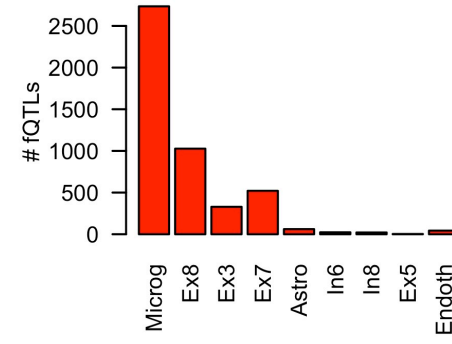
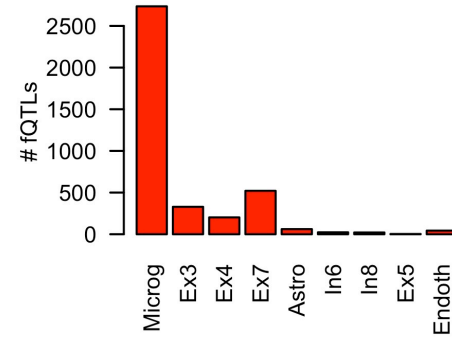
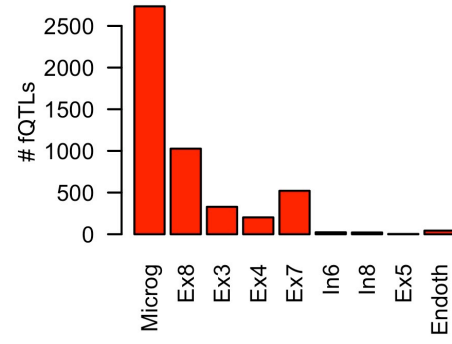
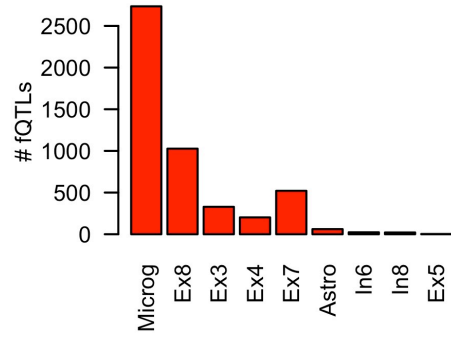
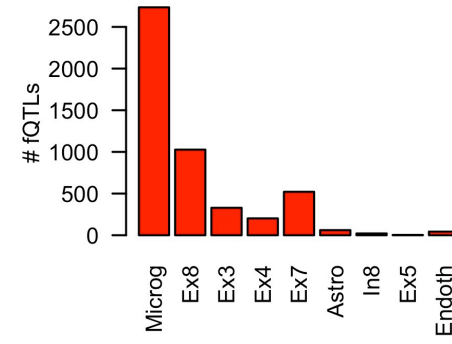
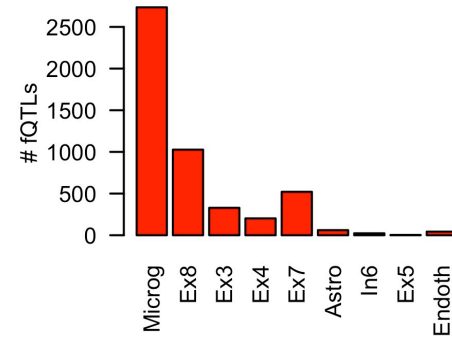
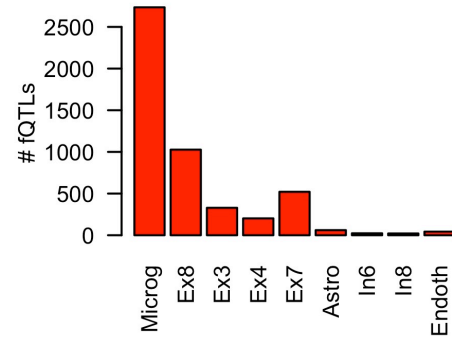
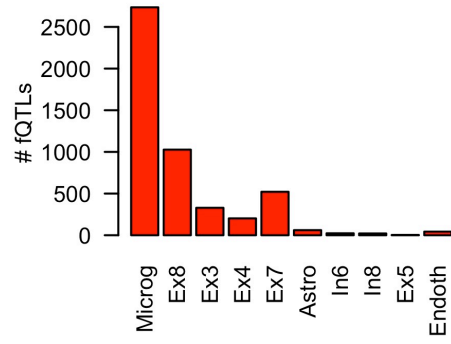
# fQTLs per cell type for the set of size 12



# fQTLs per cell type for the sets of size 11

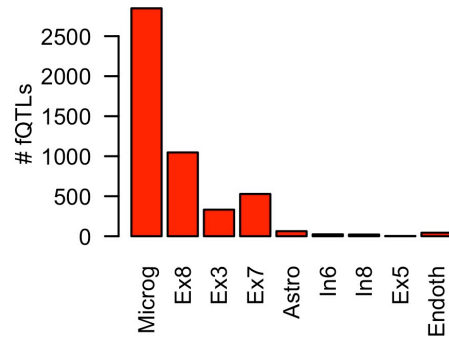
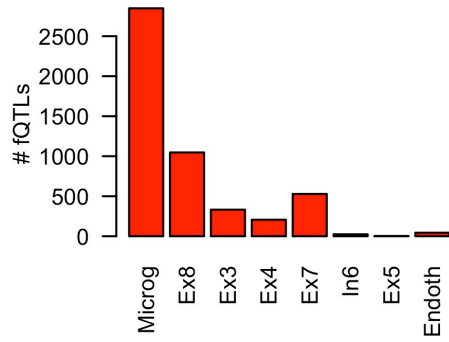
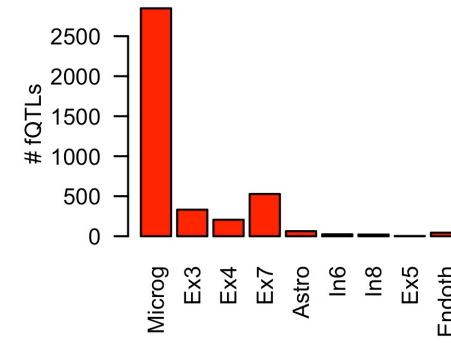
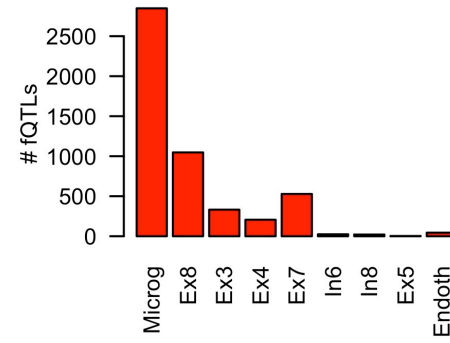
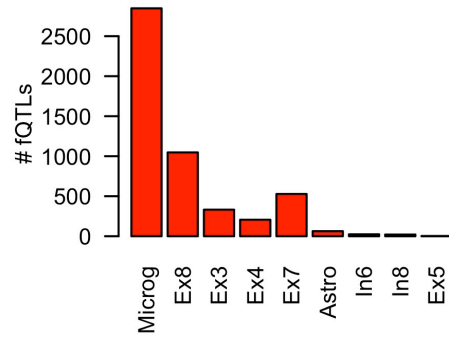
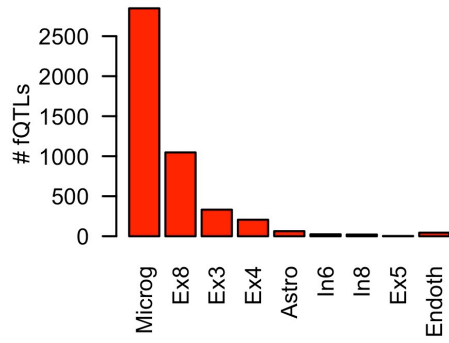
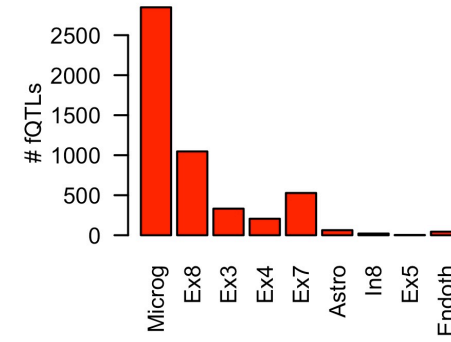
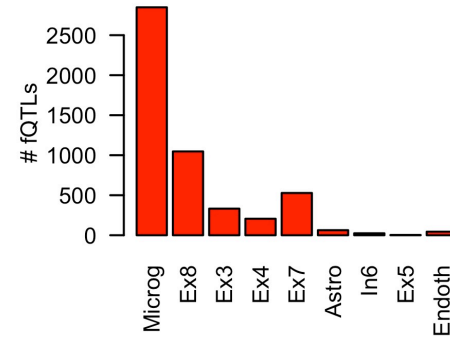
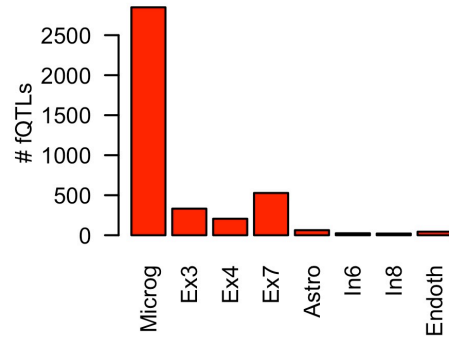
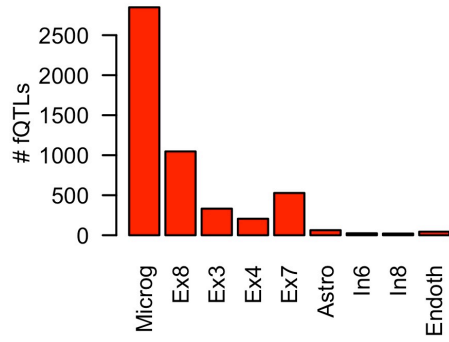


# fQTLs per cell type for the sets of size 10

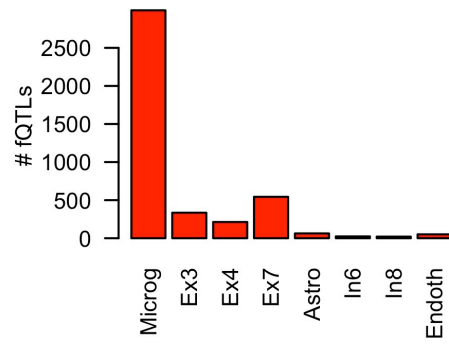
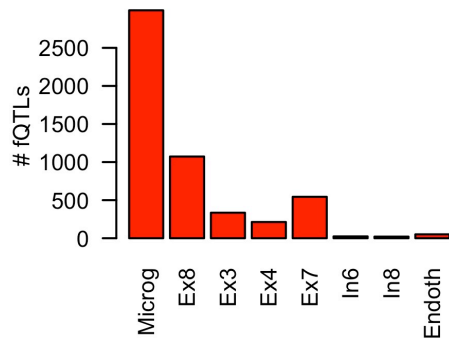
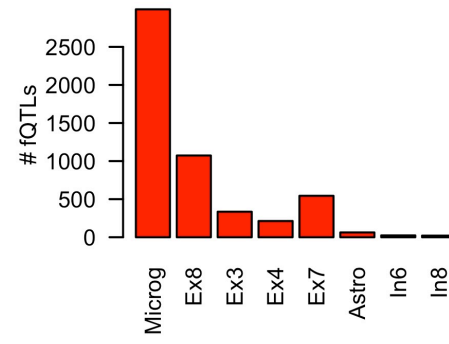
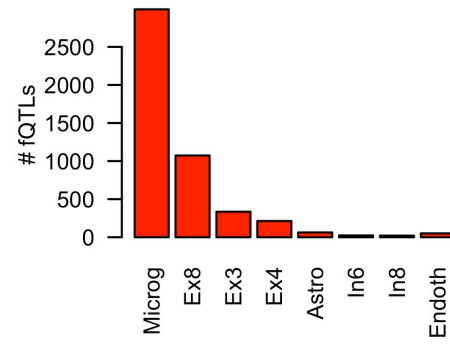
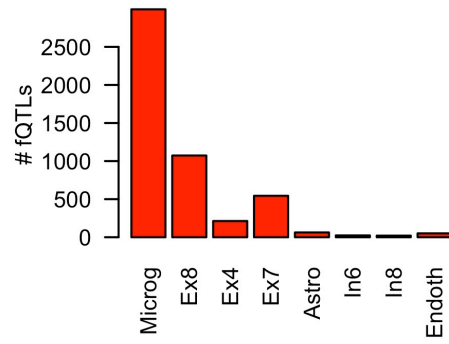
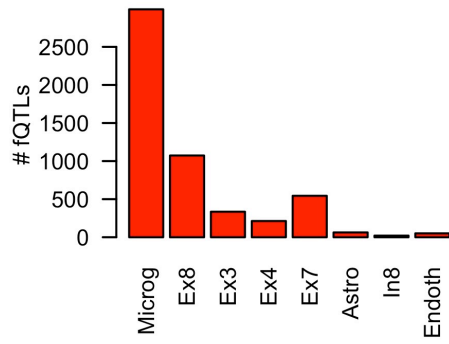
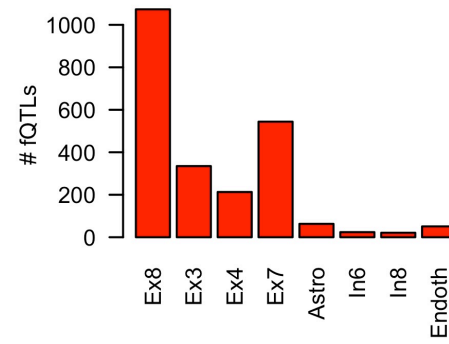
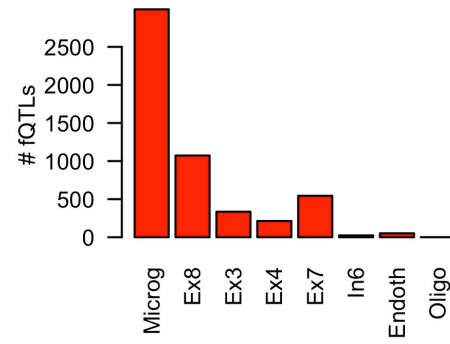
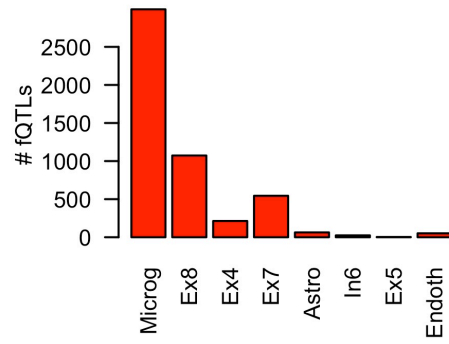
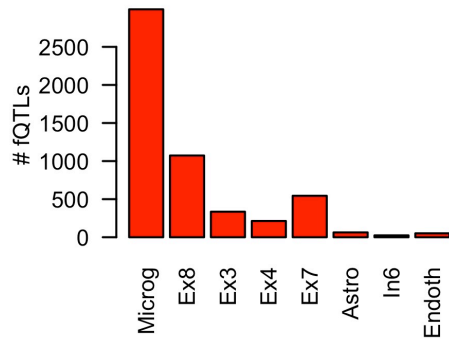




# fQTLs per cell type for the sets of size 9



# fQTLs per cell type for the sets of size 8



# fQTLs per cell type for the sets of size 7

