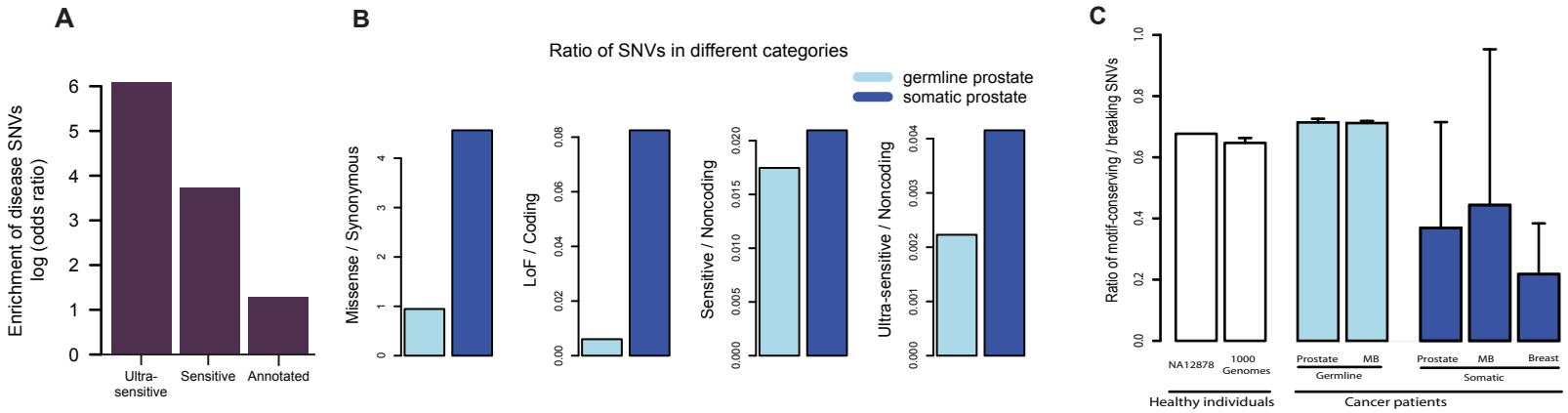
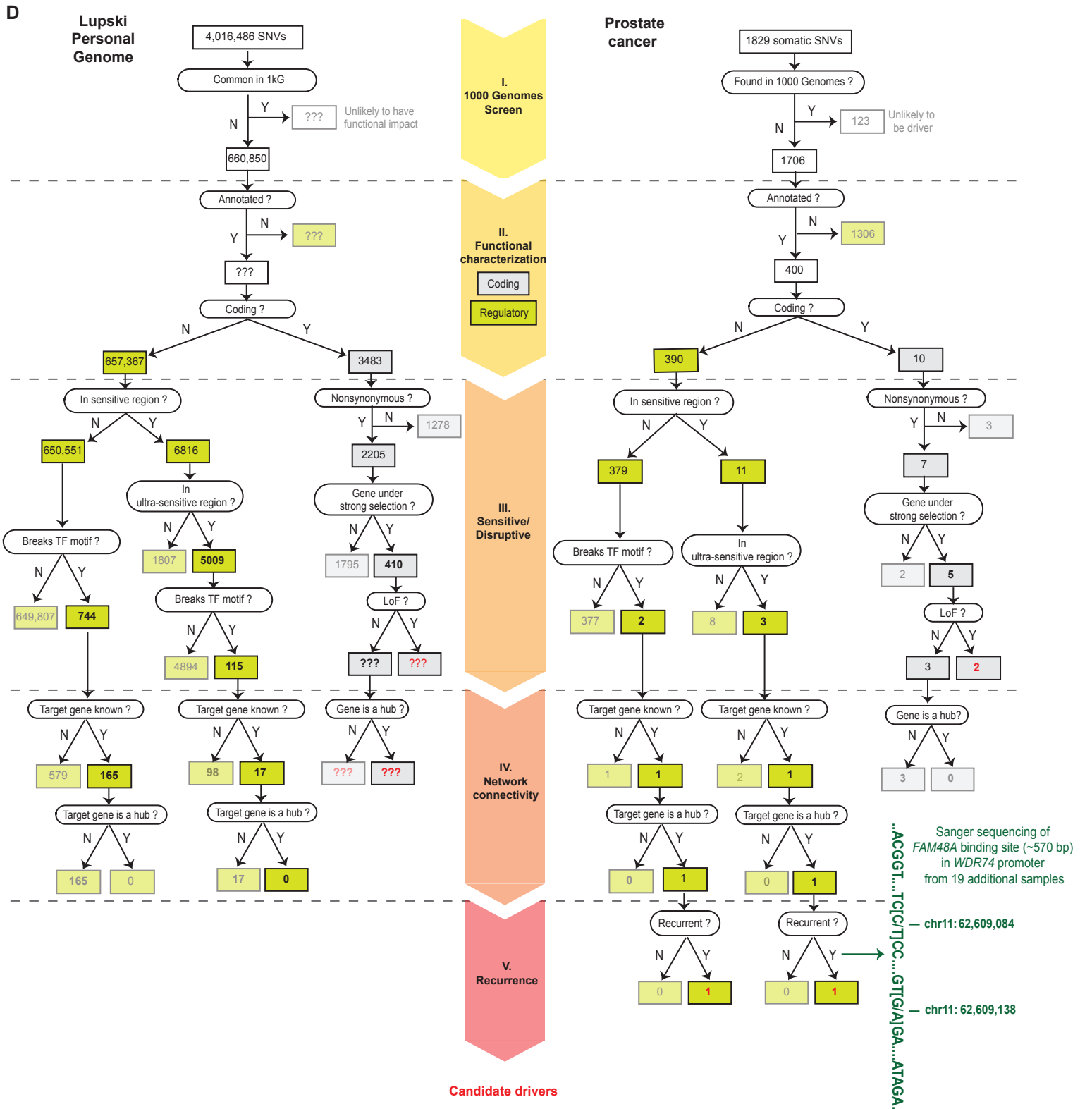
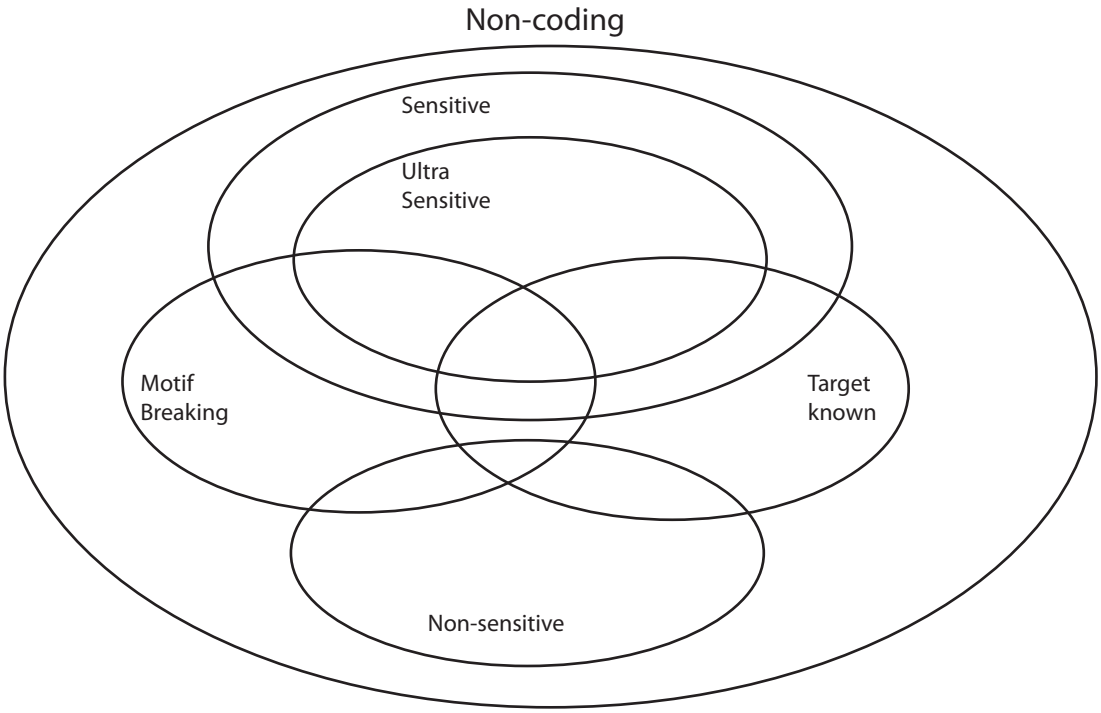
Figure 6


A
Enrichment of disease SNVs
log (odds ratio)
Ultra-sensitive
Sensitive
Annotated
B
Ratio of SNVs in different categories
germline prostate
somatic prostate
Missense / Synonymous
LoF / Coding
Sensitive / Noncoding
Ultra-sensitive / Noncoding
C
Ratio of motif-conserving / breaking SNVs
NA12878
1000 Genomes
Prostate
MB
Germline
Prostate
MB
Breast
Somatic
Healthy individuals
Cancer patients


D
Lupski Personal Genome
4,016,486 SNVs
Common in 1kG
???
Unlikely to have functional impact
660,850
Annotated ?
???
Coding ?
657,367
3483
In sensitive region ?
650,551
6816
Breaks TF motif ?
649,807
744
In ultra-sensitive region ?
1807
5009
Breaks TF motif ?
4894
115
Nonsynonymous ?
1278
2205
Gene under strong selection ?
1795
410
LoF ?
???
???
Target gene known ?
579
165
Target gene is a hub ?
165
0
Target gene known ?
98
17
Target gene is a hub ?
17
0
Gene is a hub ?
???
???
I. 1000 Genomes Screen
II. Functional characterization
Coding
Regulatory
III. Sensitive/ Disruptive
IV. Network connectivity
V. Recurrence
Candidate drivers
Prostate cancer
1829 somatic SNVs
Found in 1000 Genomes ?
123
Unlikely to be driver
1706
Annotated ?
1306
400
Coding ?
390
10
In sensitive region ?
379
11
Breaks TF motif ?
377
2
In ultra-sensitive region ?
8
3
Nonsynonymous ?
3
7
Gene under strong selection ?
2
5
LoF ?
3
2
Target gene known ?
1
1
Target gene is a hub ?
0
1
Target gene known ?
2
1
Target gene is a hub ?
0
1
Gene is a hub?
3
0
Recurrent ?
0
1
Recurrent ?
0
1
Sanger sequencing of FAM48A binding site (~570 bp) in WDR74 promoter from 19 additional samples
..ACGGT....TC[C/T]CC....GT[G/A]GA....ATAGA..
chr11: 62,609,084
chr11: 62,609,138


Non-coding
Sensitive
Ultra
Sensitive
Motif
Breaking
Target
known
Non-sensitive