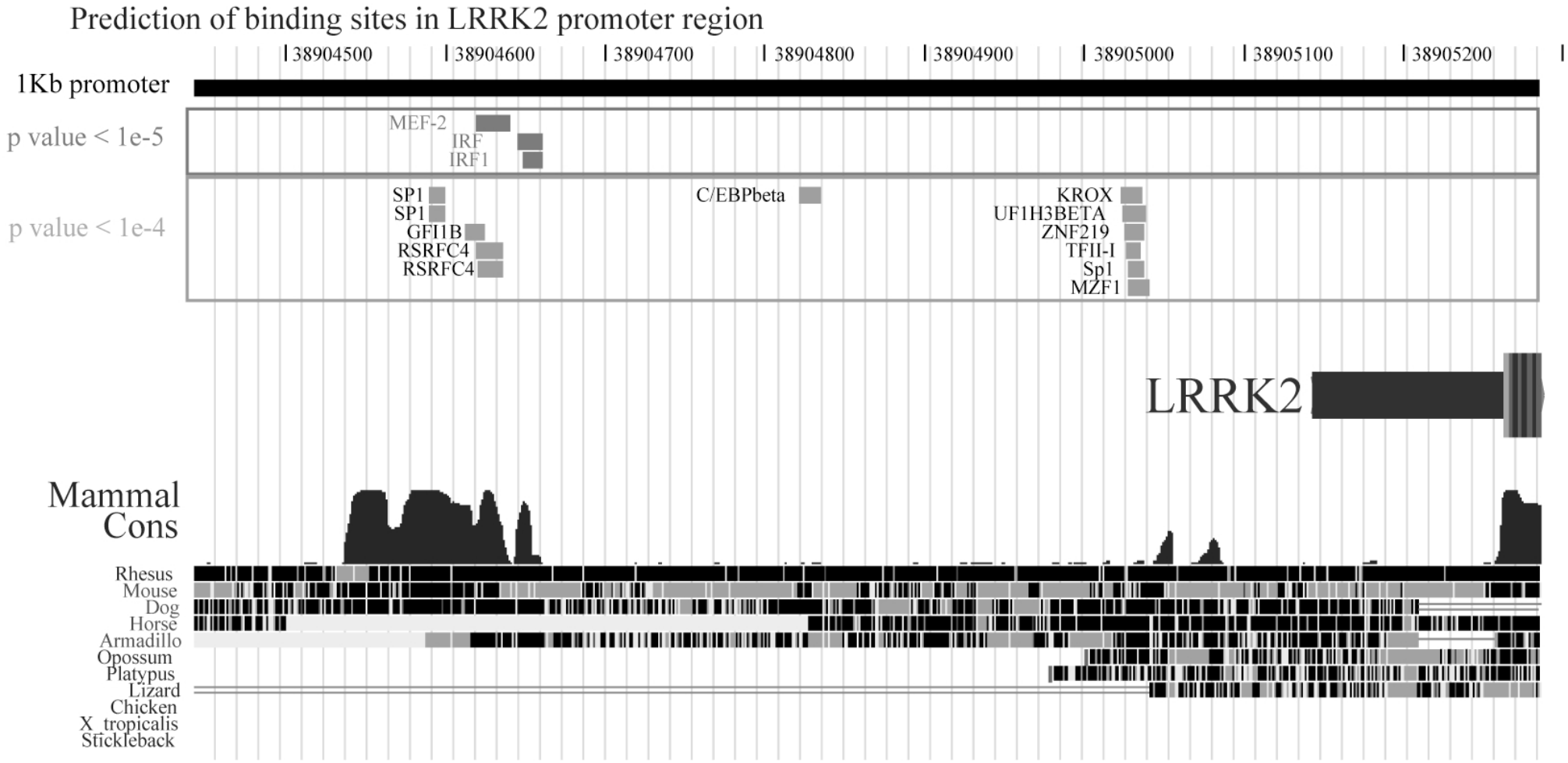
Supplementary figure 1: LRRK2 transcription factor binding site prediction.

The CREAD package was employed to identify transcription factor binding sites in 1 Kb proximal promoters of all known human genes. Known genes, their genomic location, and 28 species conservation were obtained from UCSC Genome Browser. For each protein-coding gene, the proximal promoter was defined as 700 bases upstream and 300 bases downstream of the transcription start site. Position weight matrices (PWMs) were obtained from TRANSFAC (v8.2) and JASPAR (v3.0) and each matrix was mapped to a transcription factor. We employed the program STORM for matching PWMs to the promoter sequences with $p < 10e-4$ and $p<10e-5$. Next we employed multiSTORM and accepted all sites identified in at least four species with the same parameter used in STORM. Finally, the program site-cons was employed to select for sites significantly ($p<0.05$) more conserved compared to the 100 flanking bases. Image shown was generated using custom tracks in the UCSC genome browser.

Sup figure 1


Prediction of binding sites in LRRK2 promoter region
38904500
38904600
38904700
38904800
38904900
38905000
38905100
38905200
1Kb promoter
p value < 1e-5
MEF-2
IRF
IRF1
p value < 1e-4
SP1
SP1
GFI1B
RSRFC4
RSRFC4
C/EBPbeta
KROX
UF1H3BETA
ZNF219
TFII-I
Sp1
MZF1
LRRK2
Mammal Cons
Rhesus
Mouse
Dog
Horse
Armadillo
Opossum
Platypus
Lizard
Chicken
X_tropicalis
Stickleback