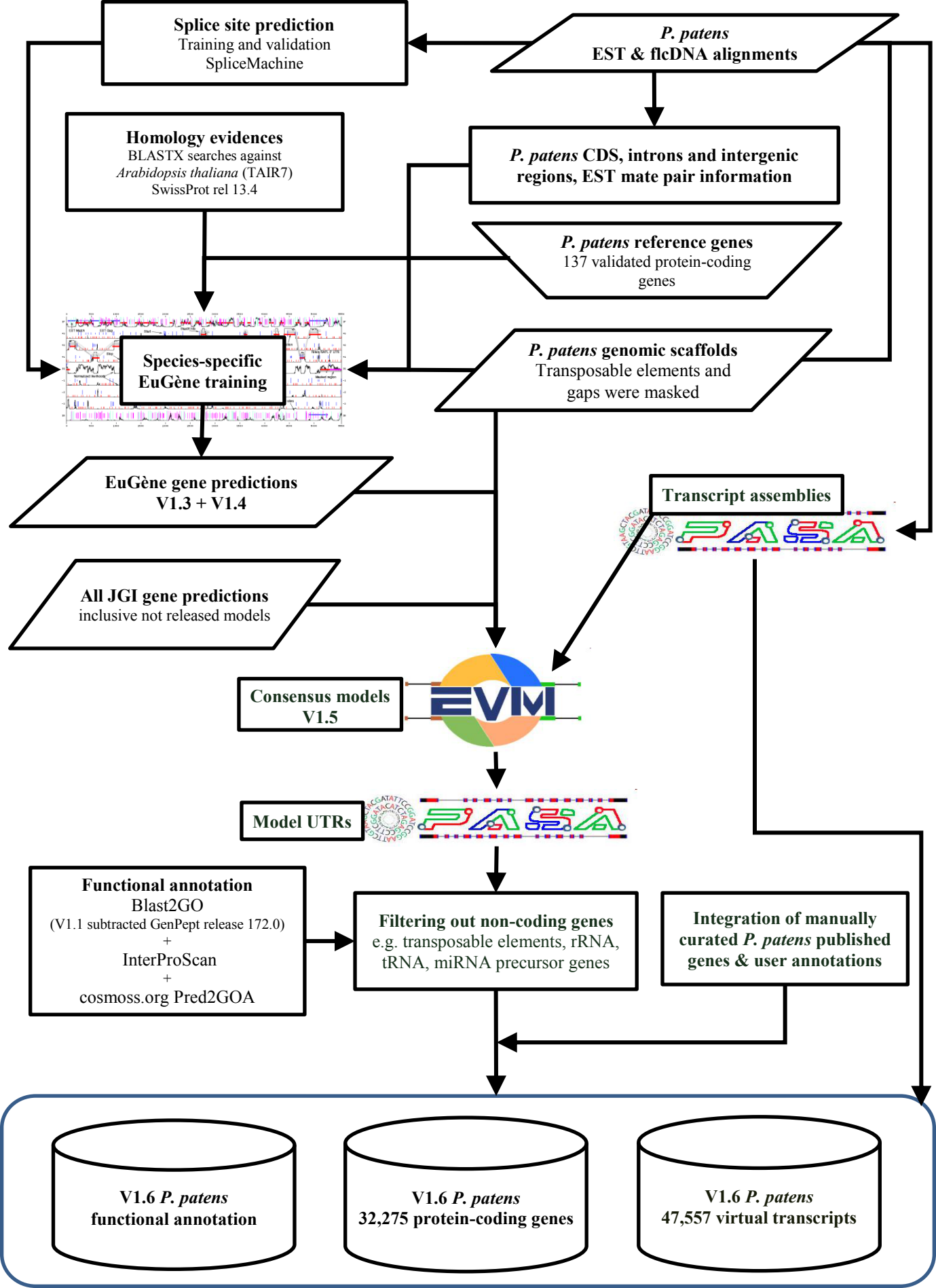
Splice site prediction
Training and validation
SpliceMachine
P. patens
EST & flcDNA alignments
Homology evidences
BLASTX searches against
Arabidopsis thaliana (TAIR7)
SwissProt rel 13.4
P. patens CDS, introns and intergenic regions, EST mate pair information
P. patens reference genes
137 validated protein-coding genes
Species-specific EuGène training
P. patens genomic scaffolds
Transposable elements and gaps were masked
EuGène gene predictions
V1.3 + V1.4
Transcript assemblies
PASA
All JGI gene predictions
inclusive not released models
Consensus models
V1.5
EVM
Model UTRs
PASA
Functional annotation
Blast2GO
(V1.1 subtracted GenPept release 172.0)
+
InterProScan
+
cosmoss.org Pred2GOA
Filtering out non-coding genes
e.g. transposable elements, rRNA, tRNA, miRNA precursor genes
Integration of manually curated P. patens published genes & user annotations
V1.6 P. patens
functional annotation
V1.6 P. patens
32,275 protein-coding genes
V1.6 P. patens
47,557 virtual transcripts