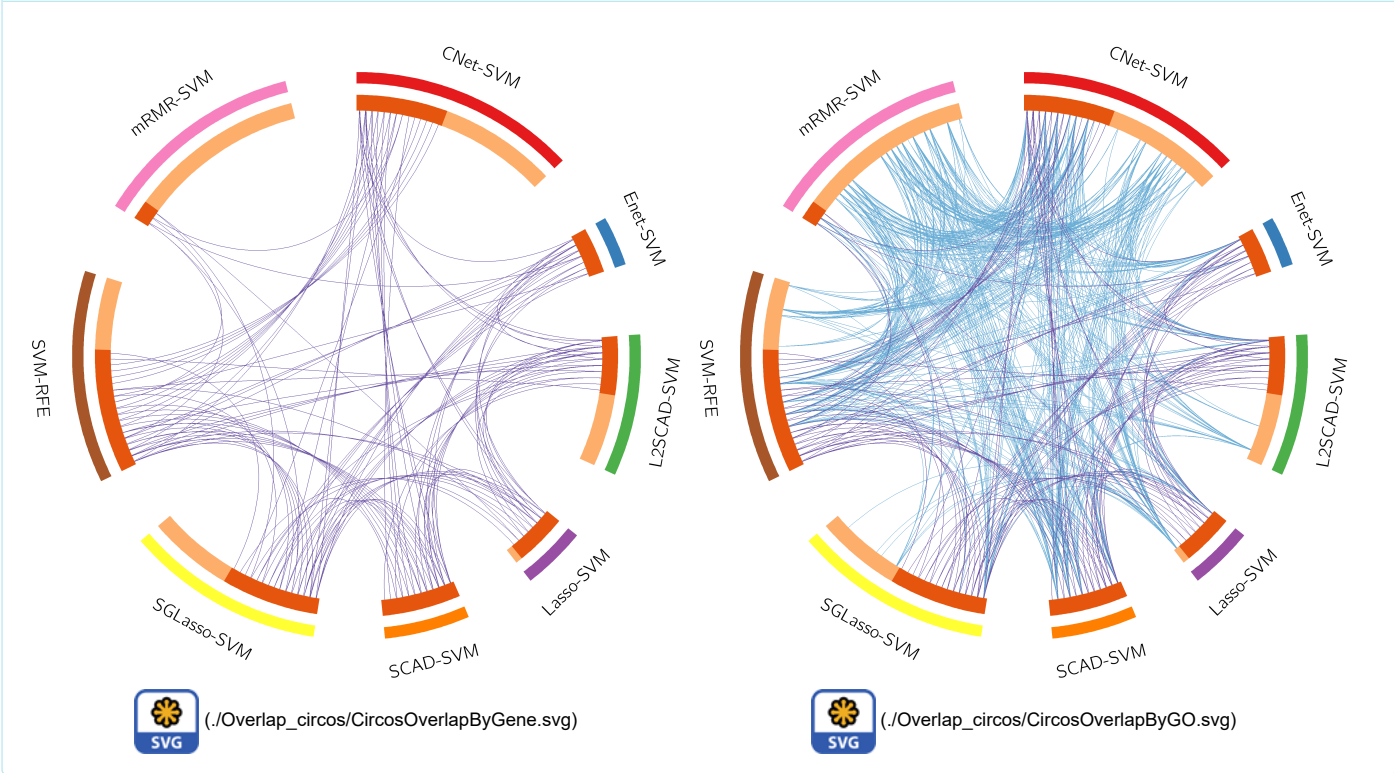


Name	Total	Unique	Color Code
SGLasso-SVM	28	28	
SVM-RFE	30	30	
mRMR-SVM	30	30	

The overlaps between these lists are shown in a Circos (<http://circo.ca>)<sup>3</sup> plot (Figure 2.a). Another useful representation is to overlap genes based on their functions or shared pathways. The overlaps between gene lists can be significantly improved by considering overlaps between genes sharing the same enriched ontology term(s) (Figure 2.b). Only ontology terms that contain less than 100 genes are used to calculate functional overlaps to avoid linking genes using very general annotation. (We do not want to link all genes, only genes that belong to specific biological processes.)

Figure 2. Overlap between gene lists: (a) only at the gene level, where purple curves link identical genes; (b) including the shared term level, where blue curves link genes that belong to the same enriched ontology term. The inner circle represents gene lists, where hits are arranged along the arc. Genes that hit multiple lists are colored in dark orange, and genes unique to a list are shown in light orange. The publication-quality version of the figures is included in the Zip package as a .svg file under the Overlap\_circos folder (readable by popular web browsers and Adobe Illustrator).



## Gene Annotation

The following are the list of annotations retrieved from the latest version of the database (last updated on 2022-04-22) (Table 2).

Table 2. Gene annotations extracted		
Name	Type	Description
Gene Symbol	Description	Primary HUGO gene symbol.
Description	Description	Short description.
Biological Process (GO)	Function/Location	Descriptions summarized based on gene ontology database, where up to three most informative GO terms are kept.
Kinase Class (UniProt)	Function/Location	Detailed kinase classes.
Protein Function (Protein Atlas)	Function/Location	Protein Function (Protein Atlas)
Subcellular Location (Protein Atlas)	Function/Location	Sucellular Location (Protein Atlas)
Drug (DrugBank)	Genotype/Phenotype/Disease	Drug information for the given gene as target.
Canonical Pathways	Ontology	Canonical Pathways
Hallmark Gene Sets	Ontology	Hallmark Gene Sets