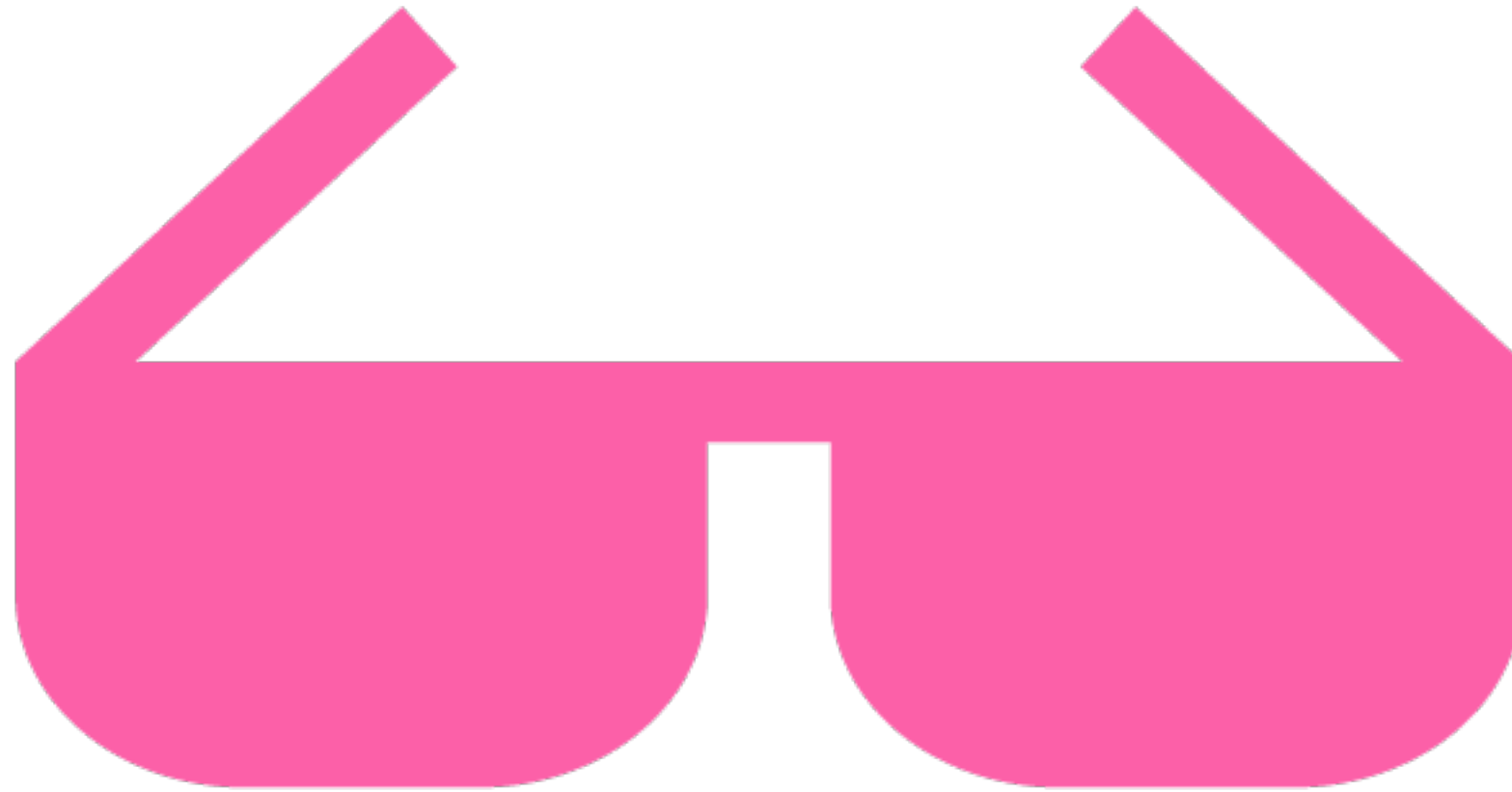


**AWESOME team**



**awesome**

# One project idea...topic modeling

GitHub, Inc. [US] | <https://github.com/NCBI-Hackathons/clusterduck>

For quick access, place your bookmarks here on the bookmarks bar. [Import bookmarks now...](#)

README.md

## ClusterDuck

Disease clustering from phenotypic literature data through Document Understanding, Comprehension and Knowledge

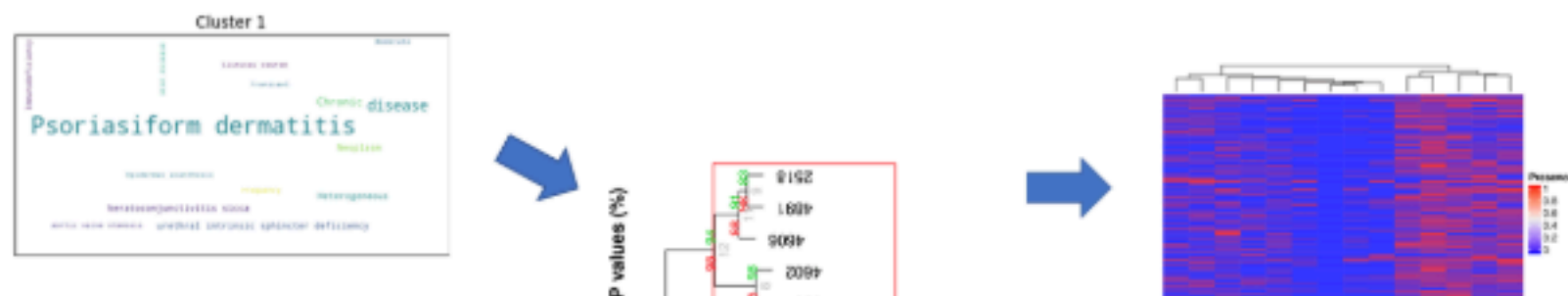
(doi: )

**PhenoX**  
phenotypeXpression

researchers and clinicians  
NPs.

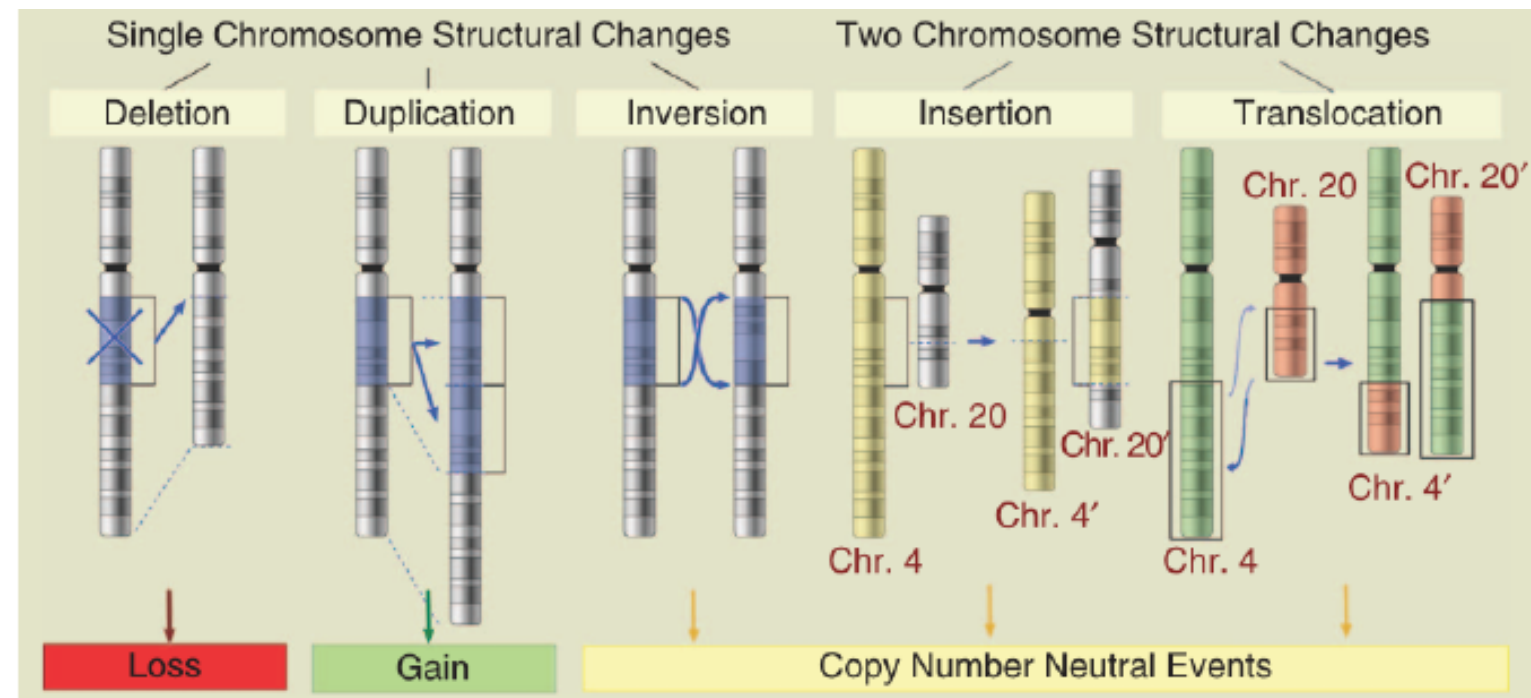
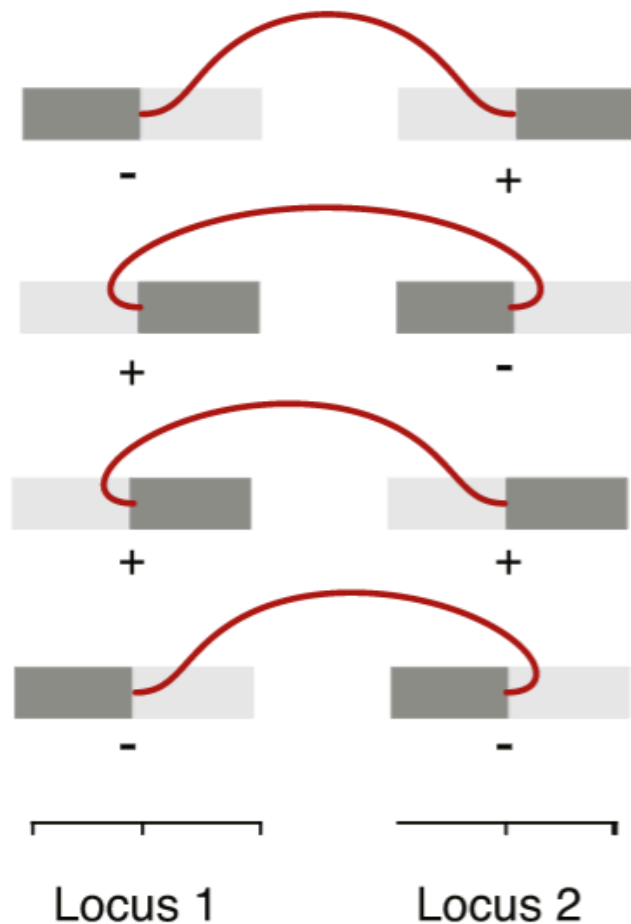
### Subclassification of disease states based on the intersection of literature and expression

PhenotypeXpression (PhenoX) is a prototype precision medicine tool for clinicians to rapidly aggregate the publicly available gene expression data and literature available from the NCBI. It requires a minimum of resources and can quickly survey publicly available research. By combining the subtype HPO and DOID information with GEO dataset expression profiles, each sub-classification of a disease is given a characteristic fingerprint. This enables clinicians to (1) identify a set of genes that have altered expression in a phenotype and its related targets to aid in research, (2) use those target genes to inform treatment selection by affected pathway, and (3) rapidly identify public datasets that are available for comparison with their own gene expression data.

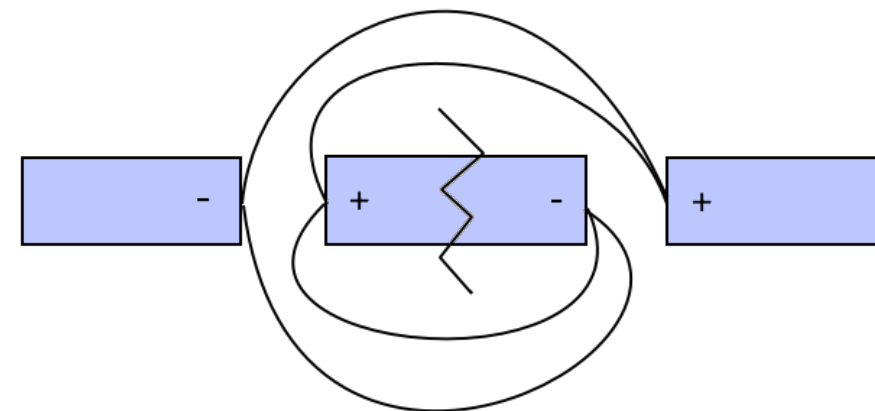


# Structural Variations

- **(-)** left-side of breakpoint
- **(+)** right-side of breakpoint



- + deletion
- + - tan. dup.
- + + inversion
- - inversion



# Goal: Build R package to annotate SV BEDPEs



**Challenges: translocations? Gene fusions? Features?**