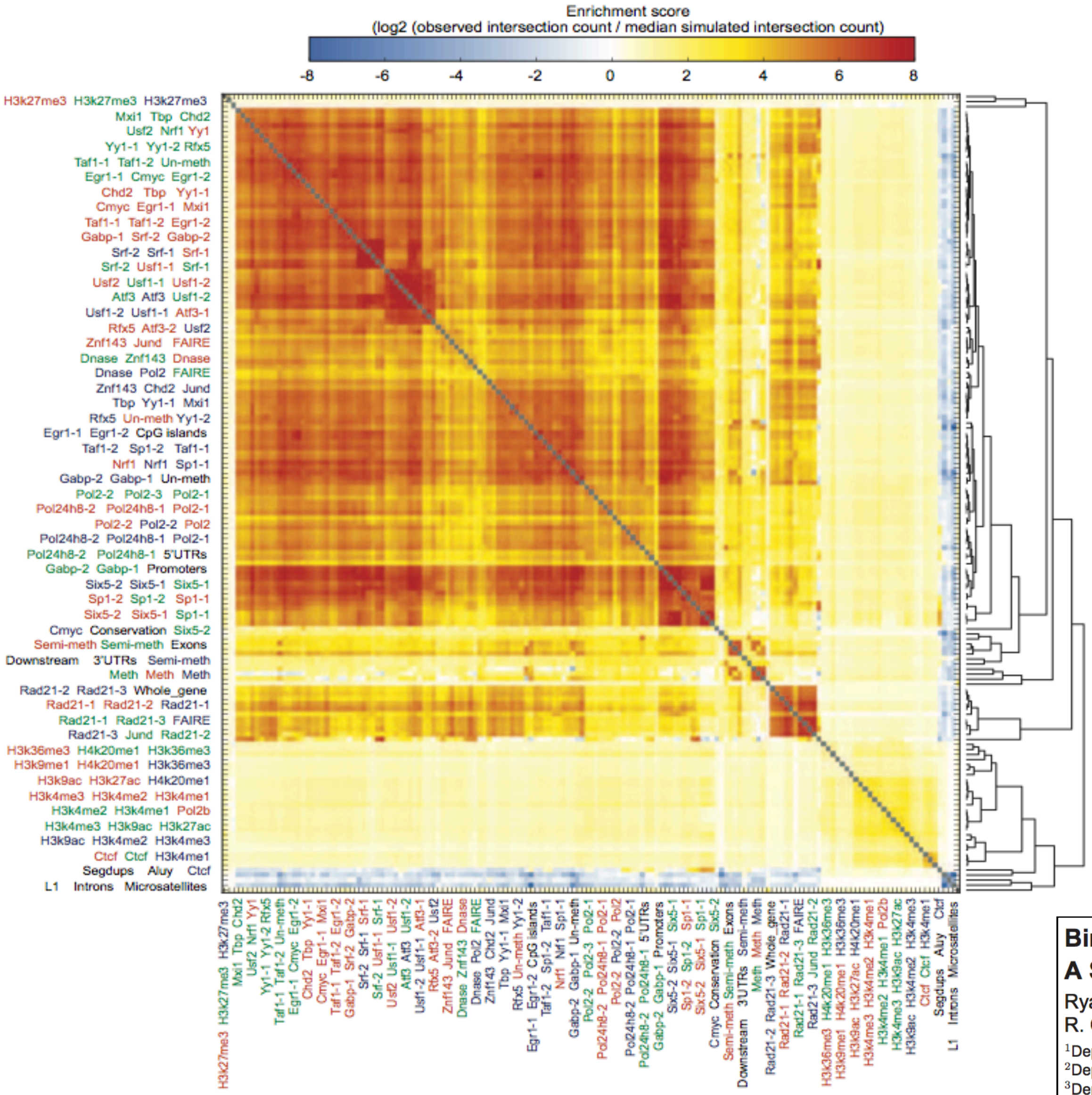
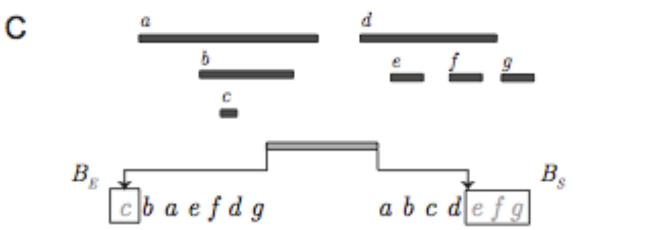
One genome. Many cell types. How?





***Step 1**: Binary search of query.**start** into database sorted by **end** coordinates. This excludes database interval c.*

***Step 2**: Binary search of query.**end** into database sorted by **start** coordinates. This excludes database intervals e,f,g.*

***Step 3**: Infer the count of intervals that intersect the query by substracting the size of the excluded set from the size of the database*

Algorithm 1: Single interval intersection counter

```
Input: Sorted interval starts and ends $B_S$ and $B_E$, query interval $a$
Output: Number of intervals $c$ intersecting $a$

Function ICOUNT($B_S, B_E, a$) begin
    $first \leftarrow$ BINARYSEARCH($B_S, a.end$)
    $last \leftarrow$ BINARYSEARCH($B_E, a.start$)
    $c \leftarrow first - last$        /* $= |B| - (last + (|B| - first))$ */
    return $c$
```

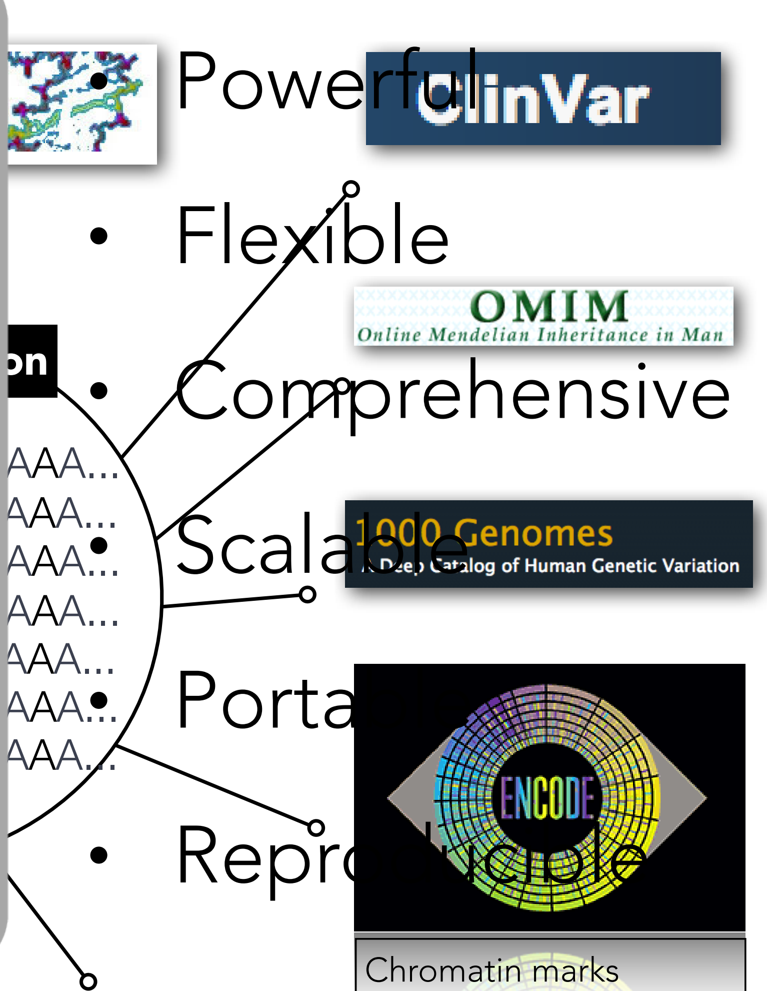
**Binary Interval Search (BITS):
A Scalable Algorithm for Counting Interval Intersections**

Ryan M. Layer[1], Kevin Skadron [1], Gabriel Robins[1], Ira M. Hall[2], and Aaron R. Quinlan[3]*

[1]Department of Computer Science, University of Virginia, Charlottesville, VA
[2]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA
[3]Department of Public Health Sciences and Center for Public Health Genomics, University of Virginia, Charlottesville, VA

Interpreting human genomes

- Powerful
- Flexible
- Comprehensive
- Scalable
- Portable
- Reproducible


ClinVar
OMIM
Online Mendelian Inheritance in Man
on
AAA...
AAA...
AAA...
AAA...
AAA...
AAA...
AAA...
1000 Genomes
A Deep Catalog of Human Genetic Variation
ENCODE

Chromatin marks
DNA methylation
RNA expression
TF binding

KEGG
Kyoto Encyclopedia of Genes and Genomes
HGMD
Human Protein Reference Database

Framework for personal, medical, and population genomics