**A novel detection method for high-order SNP epistatic interactions based on explicit-encoding-based multitasking harmony search**

## disease models

### 8 EINME models

### Table S1. Eight EINME models. The 3rd column denotes whether the model satisfies the Hardy-Weinberg equilibrium (HWE). In column 4–column 8, the values represent the prediction accuracy from k-order (k=1, 2,…,5) epistatic interaction.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | **k-order** | **HWE** | **1-order(sd)** | **2-order(sd)** | **3-order(sd)** | **4-order(sd)** | **5-order(sd)** | **tar.gz link** |
| EINME-1 | 3 | No | .502(.001) | .511(.007) | .886(.023) |  |  | [threewayBests](http://discovery.dartmouth.edu/model_free_data/evolution/threeway/threewayBests.tar.gz) |
| EINME-2 | 3 | Yes | .504(.002) | .509(.003) | .680(.024) |  |  | [HWthreewayBests](http://discovery.dartmouth.edu/model_free_data/evolution/HWthreeway/HWthreewayBests.tar.gz) |
| EINME-3 | 4 | No | .502(.001) | .510(.003) | - | .897(.018) |  | [fourwayBests](http://discovery.dartmouth.edu/model_free_data/evolution/fourway/fourwayBests.tar.gz) |
| EINME-4 | 4 | Yes | .507(.003) | .513(.003) | - | .673(.009) |  | [HWfourwayBests](http://discovery.dartmouth.edu/model_free_data/evolution/HWfourway/HWfourwayBests.tar.gz) |
| EINME-5 | 4 | No | .501(.000) | .504(.001) | .518(.003) | .567(.010) |  | [fourwayNoLowBests](http://discovery.dartmouth.edu/model_free_data/evolution/fourwayNoLow/fourwayNoLowBests.tar.gz) |
| EINME-6 | 5 | No | .502(.001) | .510(.002) | - | - | .895(.009) | [fivewayBests](http://discovery.dartmouth.edu/model_free_data/evolution/fiveway/fivewayBests.tar.gz) |
| EINME-7 | 5 | Yes | .511(.003) | .518(.003) | - | - | .693(.008) | [HWfivewayBests](http://discovery.dartmouth.edu/model_free_data/evolution/HWfiveway/HWfivewayBests.tar.gz) |
| EINME-8 | 5 | No | .503(.001) | .508(.001) | .518(.002) | .543(.004) | .690(.008) | [fivewayNoLowBests](http://discovery.dartmouth.edu/model_free_data/evolution/fivewayNoLow/fivewayNoLowBests.tar.gz) |

The eight datasets are generated by Himmelstein et al, 2011[1], which disables the discovery of disease-causing models for certain existing heuristic methods due to the lack of clues of causative SNP markers.

### 12 EIME models

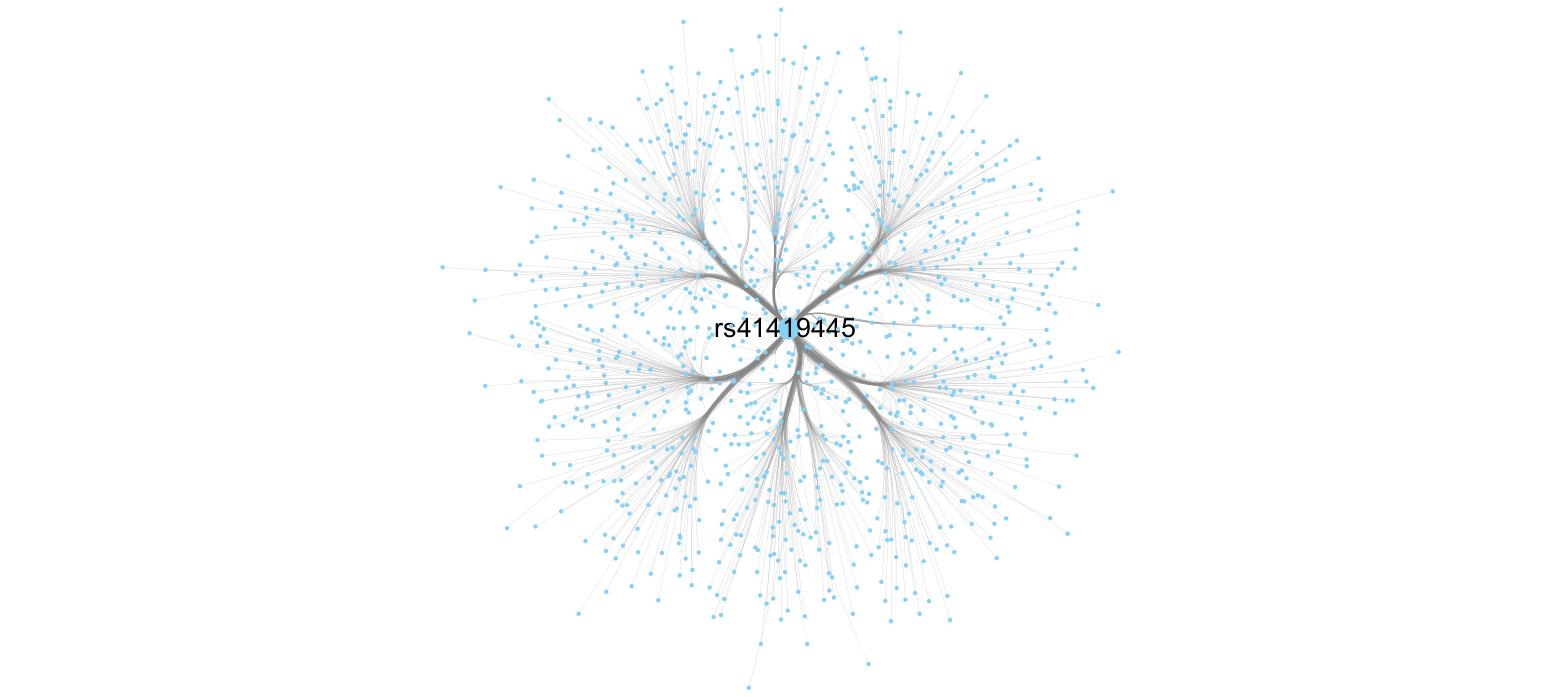
### Table S2. The parameters and the values of penetrance of 12 EIME models.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model type** | **EIME** | **order** | **Heritability(H2)** | **MAF** | **Heterogeneity proportion** |
| Additive model | EIME -1 | 5 | 0.1 | 0.1 | 1.0 |
| EIME -2 | 5 | 0.1 | 0.25 | 1.0 |
| EIME -3 | 5 | 0.1 | 0.5 | 1.0 |
| EIME -4 | 5 | 0.1 | 0.2 | 1.0 |
| Threshold model | EIME -5 | 5 | 0.1 | 0.1 | 1.0 |
| EIME -6 | 5 | 0.25 | 0.1 | 1.0 |
| EIME -7 | 5 | 0.5 | 0.1 | 1.0 |
| EIME -8 | 5 | 0.1 | 0.2 | 1.0 |
| Multiplicative  model | EIME -9 | 4 | 0.005 | 0.1 | 1.0 |
| EIME -10 | 4 | 0.005 | 0.2 | 1.0 |
| EIME -11 | 4 | 0.005 | 0.4 | 1.0 |
| EIME -12 | 4 | 0.004 | 0.05 | 1.0 |

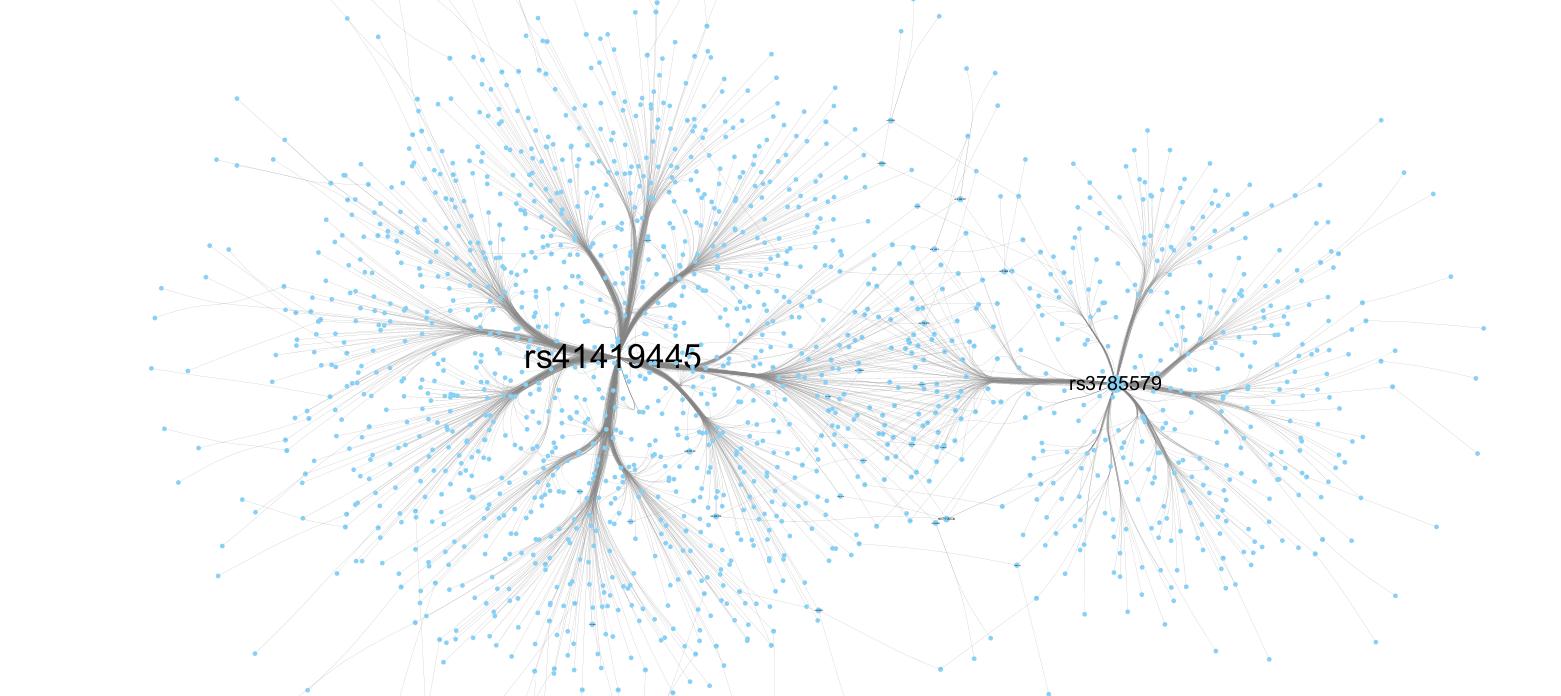
**H2** denotes the genetic heritability. **MAF** represents the minor allele frequencies.

**The datasets are generated using GAMETES software.**

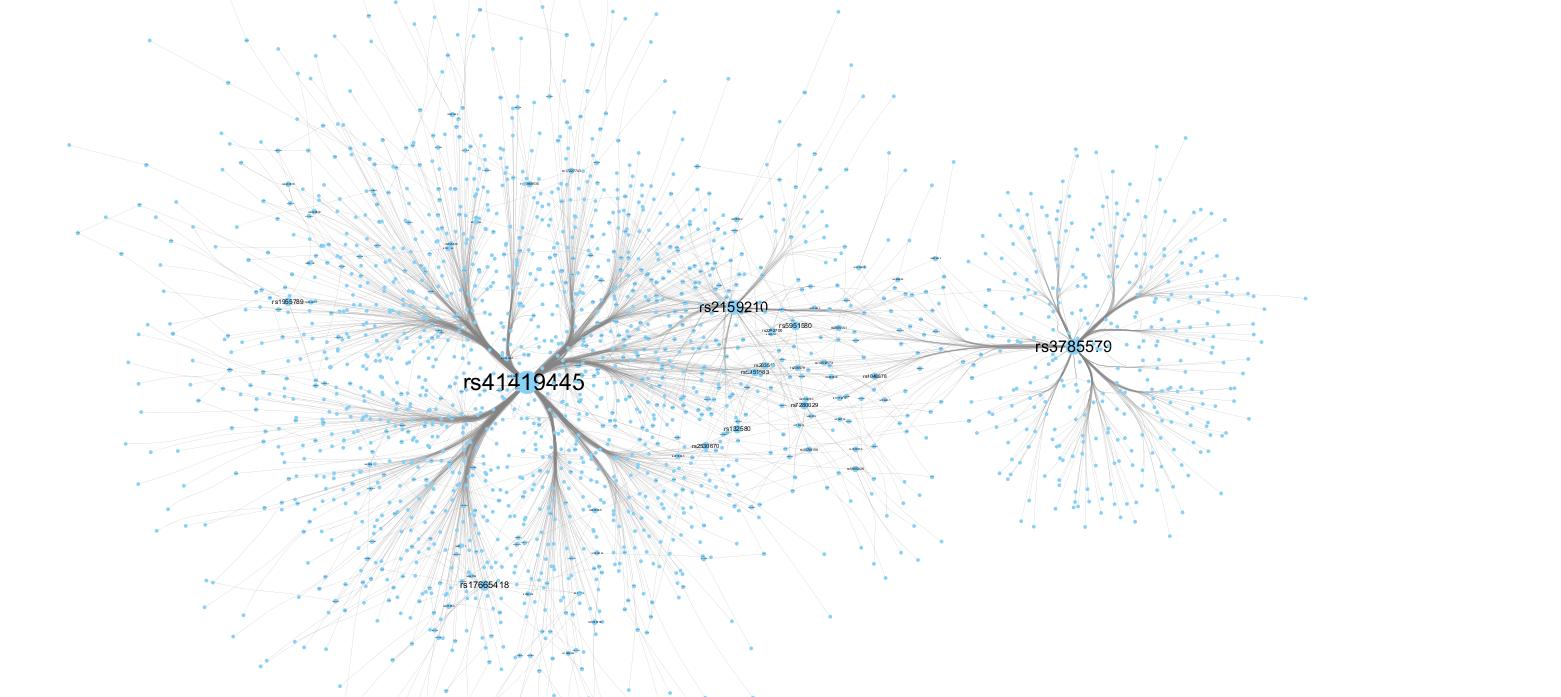
## Figures

****

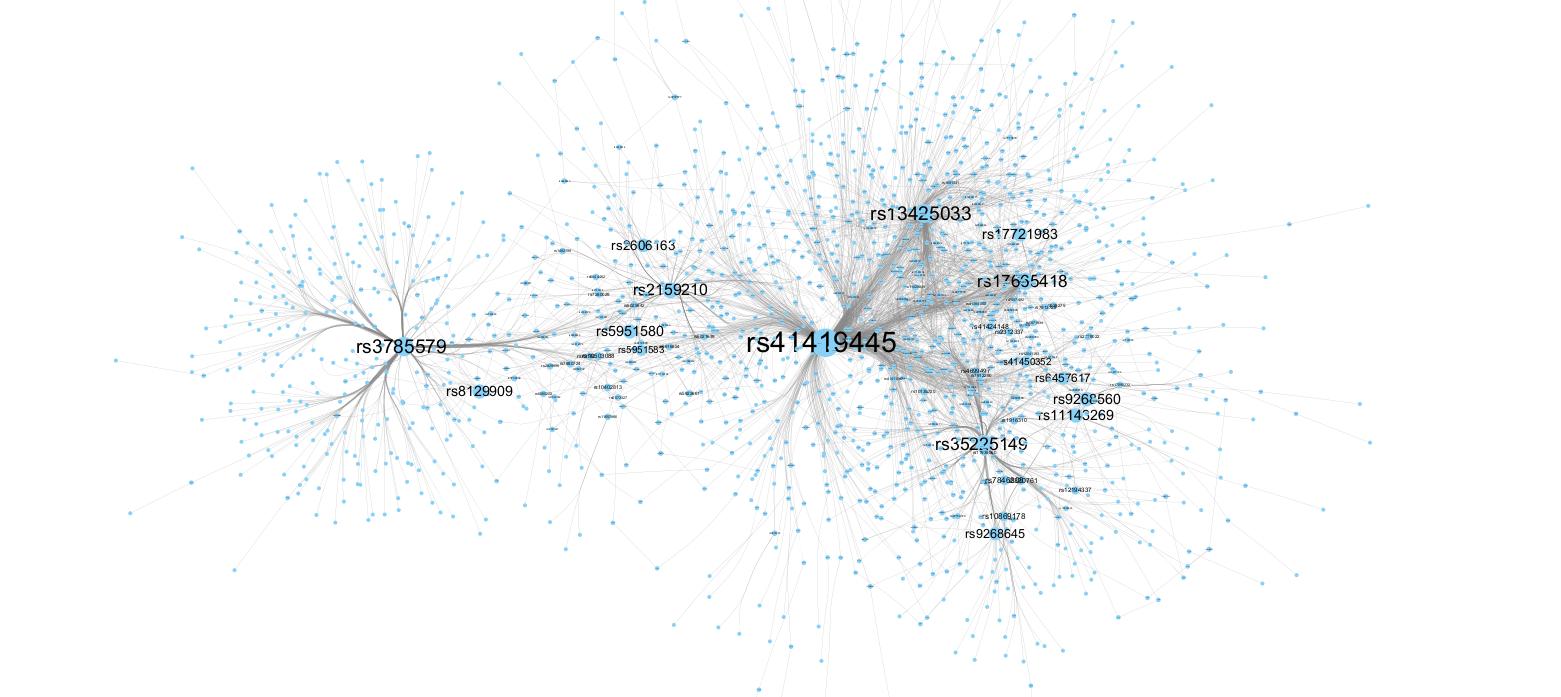
**Figure S1**. **2th-order epistatic interaction network for RA**

****

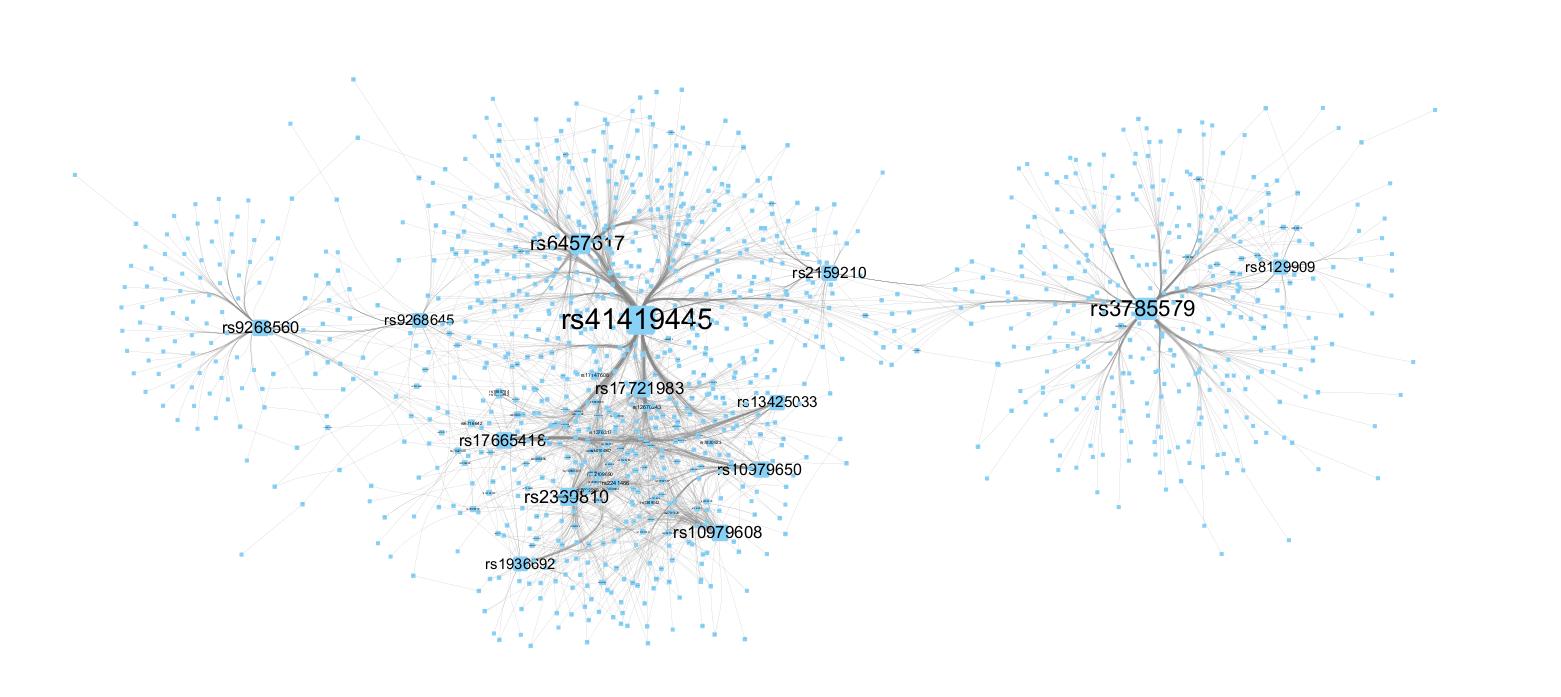
**Figure S2**. 3**th-order epistatic interaction network for RA**

****

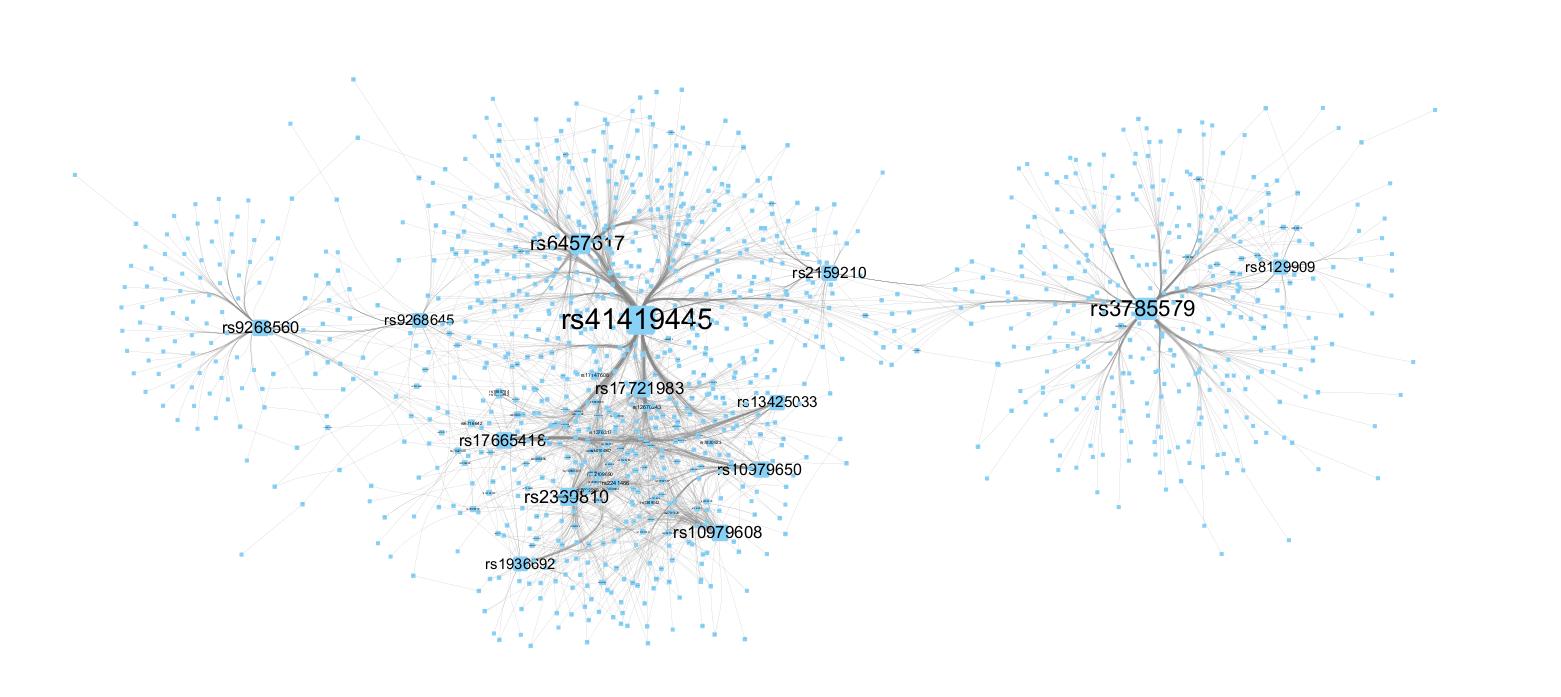
**Figure S3**. **4th-order epistatic interaction network for RA**

****

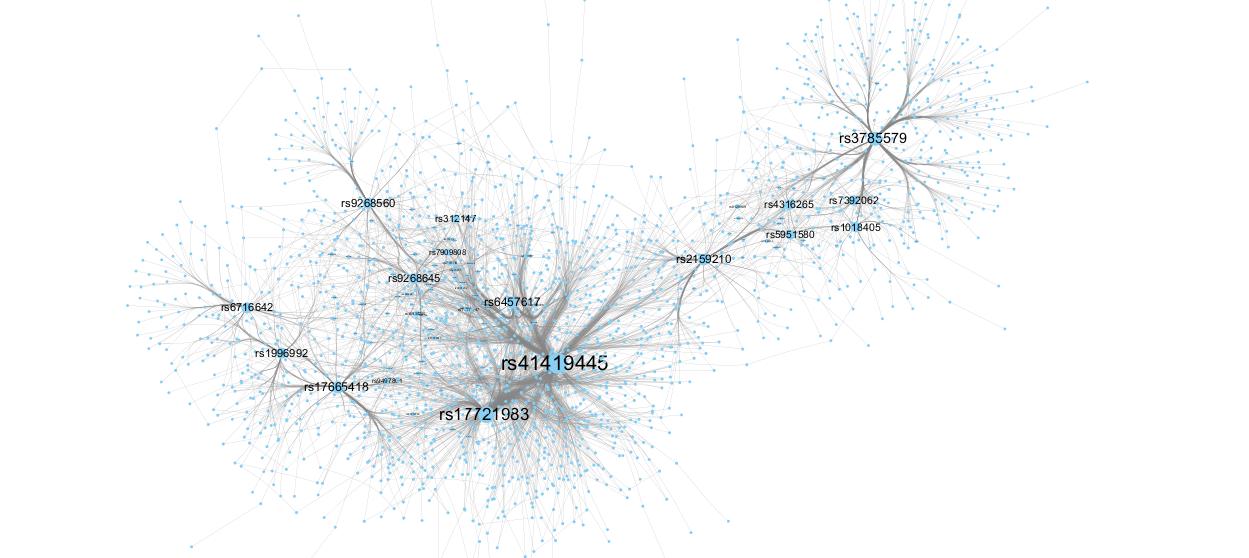
**Figure S4**. **5th-order epistatic interaction network for RA**

****

**Figure S5**. **6th-order epistatic interaction network for RA**

****

**Figure S6**. 7**th-order epistatic interaction network for RA**



**Figure S7**. **8th-order epistatic interaction network for RA**

## References

[1] Himmelstein et al. Evolving hard problems: Generating human genetics datasets with a complex etiology. BioData Mining 4, 21(2011). doi:10.1186/1756-0381-4-21. <http://discovery.dartmouth.edu/model_free_data/>.