Plink Report (with EpiGEN)

Ways of evaluating epistasis detection algorithm performance:

1. Number of interacting pairs (true positives) identified.
2. Number of false positives.
3. Position of true positive(s) (averaged if multiple) when ranked by p-value. If any true positives are missing their position is considered to be the “last” in the list.
   1. The above definition will be referred to as penalized true positive position.
   2. Alternatively, average true positive position can be computed only using the detected true positives. This will be referred to as unpenalized true positive position.
4. Runtime (~1second for 1000 SNPs, ~14 minutes for 100000 SNPs)
   1. Need to specify how the plink was executed on the HPC.
   2. Regardless Plink’s epistasis seems well optimized and may be usable on > 1000000 SNPs.

Compare these performance parameters across:

1. Multiplicative vs Dominant vs Recessive vs XOR Interaction.
2. Interaction Alpha: 1.25, 1.5, 2, 3, 8, 16.
3. Pure vs Impure.

Each EpiGEN dataset is named as follows: PURE/IMPURE (indicates presence of marginal effects) \_ MULTIPLICATIVE/DOMINANT/RECESSIVE/XOR (indicates type of interaction) \_ ONEPAIR / TWOPAIRS / EIGHTPAIRS (indicates number of interacting disease snp pairs) \_ BASELINEALPHA10 (indicates quantitative phenotype value in absence of marginal and interaction effects) \_ InteractionAlpha 1.25/1.5/2/3/8/16 (indicates strength of interaction effect)

**RESULTS (pages 1-11)**

**MULTIPLICATIVE DATASETS**

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 0

Number of false positives: 25

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 88

F1 Score: 0.022

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 174

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 49

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 555

F1 Score: 0.0036

Average true positive position (unpenalized) ranked by p-value: 30

Average true positive position (penalized) ranked by p-value: 293.5

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 185

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 213

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 29

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 151

F1 Score: 0.013

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 27

F1 Score: 0.069

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 29

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 0

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 116

F1 Score: 0.017

Average true positive position (unpenalized) ranked by p-value: 4

Average true positive position (penalized) ranked by p-value: 61

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 554

F1 Score: 0.0036

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 278.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 1

Number of false positives: 10

F1 Score: 0.15

Average true positive position (unpenalized) ranked by p-value: 3

Average true positive position (penalized) ranked by p-value: 7.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 3

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 2

Number of false positives: 787

0.0050

Average true positive position (unpenalized) ranked by p-value: 324

Average true positive position (penalized) ranked by p-value: 673.5

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 1079

0.0018

Average true positive position (unpenalized) ranked by p-value: 6

Average true positive position (penalized) ranked by p-value: 946.625

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 0

Number of false positives: 792

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 0

Number of false positives: 970

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

**DOMINANT DATASETS**

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 264

F1 Score: 0.0075

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 567

F1 Score: 0.0035

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 2

Number of false positives: 229

F1 Score: 0.017

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 2

Number of false positives: 249

F1 Score: 0.016

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 323

F1 Score: 0.0061

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 326

F1 Score: 0.0061

Average true positive position (unpenalized) ranked by p-value: 1

Average true positive position (penalized) ranked by p-value: 1

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 2

Number of false positives: 296

F1 Score: 0.013

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 2

Number of false positives: 397

F1 Score: 0.0099

Average true positive position (unpenalized) ranked by p-value: 1.5

Average true positive position (penalized) ranked by p-value: 1.5

**RECESSIVE DATASETS**

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 3

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 2

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 1502

F1 Score: 0.0013

Average true positive position (unpenalized) ranked by p-value: 57

Average true positive position (penalized) ranked by p-value: 780.5

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 1462

F1 Score: 0.0014

Average true positive position (unpenalized) ranked by p-value: 42

Average true positive position (penalized) ranked by p-value: 753

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 1833

F1 Score: 0.0011

Average true positive position (unpenalized) ranked by p-value: 191

Average true positive position (penalized) ranked by p-value: 191

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 1741

F1 Score: 0.0011

Average true positive position (unpenalized) ranked by p-value: 149

Average true positive position (penalized) ranked by p-value: 149

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 582

F1 Score: 0.0034

Average true positive position (unpenalized) ranked by p-value: 100

Average true positive position (penalized) ranked by p-value: 342

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 620

F1 Score: 0.0032

Average true positive position (unpenalized) ranked by p-value: 9

Average true positive position (penalized) ranked by p-value: 315.5

**XOR DATASETS**

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 1248

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 794

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 286

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 1058

F1 Score: 0.0

Average true positive position (unpenalized) ranked by p-value: 0

Average true positive position (penalized) ranked by p-value: 0

**DISCUSSION (Page 12-13)**

Unpenalized true positive positions by interaction type.

Multiplicative: 0,1,0,0,30,0,0,0,1,1,0,0,4,1,3,0,324,6,0,0

Dominant: 1,1,1.5,1.5,1,1,1.5,1.5

Recessive: 0,0,57,42,191,149,100,9

XOR: 0,0,0,0

F1 Score by grouped by interaction type.

Multiplicative: 0.022, 0.0036, 0.013, 0.069, 0.017, 0.0036, 0.15, 0.0050, 0.0018, 0,0,0,0,0,0,0,0,0,0,0 => 0.014

Dominant: 0.0075, 0.0035, 0.017, 0.016, 0.0061, 0.0061, 0.013, 0.0099 => 0.0099

Recessive: 0.0013, 0.0014, 0.0011, 0.0011, 0.0034, 0.0032, 0, 0 => 0.0014

XOR: 0,0,0,0 => 0

From this point onward assume all true positive positions are penalized.

Multiplicative Interaction Average True Positive Positions (note that 0 means no true positives were detected): 0,1,0,0,293.5,0,0,0,1,1,0,0,61,278.5,7.5,0,673.5,946.625,0,0

Dominant Interaction Average True Positive Positions: 1,1,1.5,1.5,1,1,1.5,1.5

Recessive Interaction Average True Positive Positions: 0,0,780.5,753,191,149,342,315.5

XOR Interaction Average True Positive Positions: 0,0,0,0

Plink’s epistasis could not detect XOR interaction. Multiplicative and recessive interactions were frequently detected.

Pairs Detected (# of SNP pairs correctly detected / total # of interacting disease SNP pairs) – only count if at least one pair was detected: 1/1,1/2,1/1,1/1,1/2,1/2,1/2,2/8,1/8, 1/1,1/1,2/2,2/2,1/1,1/1,2/2,2/2,1/2,1/2,1/1,1/1,1/2,1/2

The number of detected pairs was usually 1-2.

Now let us examine when Plink’s epistasis could detect multiplicative interaction.

Multiplicative Interaction Average True Positive Position Clustered by Interaction Alpha:

Interaction Alpha 1.25: 0,0,0,0

Interaction Alpha 1.5: 0,0,0,7.5

Interaction Alpha 2.0: 0,293.5,1,61,946.625,0

Interaction Alpha 3.0: 1,0,1,278.5,673.5,0

Clear trend here that interactions with larger effect were more easily detectable.

Let us also compare average true positive position for Pure vs Impure Epistasis in case of multiplicative interaction only.

Multiplicative Interaction Average True Positive Position Clustered by Pure Status:

Pure: 1,1,0,0,61,278.5,7.5,0,673.5,946.625

Impure: 0,1,0,0,293.5,0,0,0,0,0

Pure epistasis is easier to detect as expected.

Notably, all multiplicative pure interactions with alpha >= 2 were detected.

Notably, Plink’s epistasis rarely found more than one pair even when multiple pairs were present.