MDR 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, NA for 100000 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: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 2

Average true positive position (penalized) ranked by bal accuracy: 2

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 3

Average true positive position (penalized) ranked by bal accuracy: 3

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 2

Number of false positives: 498

F1 Score: 0.0080

Average true positive position (unpenalized) ranked by bal accuracy: 8.5

Average true positive position (penalized) ranked by bal accuracy: 8.5

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 251

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Multiplicative\_OnePair\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 251

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 2

Number of false positives: 498

F1 Score: 0.0080

Average true positive position (unpenalized) ranked by bal accuracy: 7.5

Average true positive position (penalized) ranked by bal accuracy: 7.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha15

Number of true positives: 2

Number of false positives: 498

F1 Score: 0.0080

Average true positive position (unpenalized) ranked by bal accuracy: 2.5

Average true positive position (penalized) ranked by bal accuracy: 2.5

Pure\_Multiplicative\_TwoPairs\_BaselineAlpha10\_InteractionAlpha125

Number of true positives: 2

Number of false positives: 498

F1 Score: 0.0080

Average true positive position (unpenalized) ranked by bal accuracy: 3.5

Average true positive position (penalized) ranked by bal accuracy: 3.5

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 6

Number of false positives: 494

F1 Score: 0.024

Average true positive position (unpenalized) ranked by bal accuracy: 55

Average true positive position (penalized) ranked by bal accuracy: 166.5

Pure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 5

Number of false positives: 495

F1 Score: 0.020

Average true positive position (unpenalized) ranked by bal accuracy: 98.6

Average true positive position (penalized) ranked by bal accuracy: 249.5

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha3

Number of true positives: 2

Number of false positives: 498

F1 Score: 0.0079

Average true positive position (unpenalized) ranked by bal accuracy: 10

Average true positive position (penalized) ranked by bal accuracy: 378.25

Impure\_Multiplicative\_EightPairs\_BaselineAlpha10\_InteractionAlpha2

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 23

Average true positive position (penalized) ranked by bal accuracy: 441.25

**DOMINANT DATASETS**

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 181

Average true positive position (penalized) ranked by bal accuracy: 341

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Dominant\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 1

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 7

Average true positive position (penalized) ranked by bal accuracy: 254

Pure\_Dominant\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 451

Average true positive position (penalized) ranked by bal accuracy: 476

**RECESSIVE DATASETS**

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Impure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 1

Average true positive position (penalized) ranked by bal accuracy: 251

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 70

Average true positive position (penalized) ranked by bal accuracy: 70

Pure\_Recessive\_OnePair\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 131

Average true positive position (penalized) ranked by bal accuracy: 131

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 0

Number of false positives: 500

F1 Score: 0.0

Average true positive position (unpenalized) ranked by bal accuracy: 0

Average true positive position (penalized) ranked by bal accuracy: 0

Pure\_Recessive\_TwoPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 1

Number of false positives: 499

F1 Score: 0.0040

Average true positive position (unpenalized) ranked by bal accuracy: 390

Average true positive position (penalized) ranked by bal accuracy: 445.5

**XOR DATASETS:**

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 8

Number of false positives: 492

F1 Score: 0.031

Average true positive position (unpenalized) ranked by bal accuracy: 63.75

Average true positive position (penalized) ranked by bal accuracy: 63.75

Pure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 7

Number of false positives: 493

F1 Score: 0.027

Average true positive position (unpenalized) ranked by bal accuracy: 121.71428571428571

Average true positive position (penalized) ranked by bal accuracy: 169.125

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha16

Number of true positives: 6

Number of false positives: 494

F1 Score: 0.024

Average true positive position (unpenalized) ranked by bal accuracy: 16.333333333333332

Average true positive position (penalized) ranked by bal accuracy: 137.5

Impure\_XOR\_EightPairs\_BaselineAlpha10\_InteractionAlpha8

Number of true positives: 6

Number of false positives: 494

F1 Score: 0.024

Average true positive position (unpenalized) ranked by bal accuracy: 85

Average true positive position (penalized) ranked by bal accuracy: 189

**DISCUSSION (Page 12-13)**

Unpenalized true positive positions by interaction type.

Multiplicative: 2, 3, 0, 0, 8.5, 1, 0, 0, 1, 1, 1, 1, 1, 7.5, 2.5, 3.5, 55, 98.6, 10, 23

Dominant: 0, 1, 0, 181, 1, 1, 7, 451

Recessive: 0, 0, 0, 1, 70, 131, 0, 390

XOR: 63.75, 121.71, 16.3, 85

F1 Score by grouped by interaction type.

Multiplicative: 0.004, 0.004, 0, 0, 0.008, 0.004, 0, 0, 0.004, 0.004, 0.004, 0.004, 0.004, 0.008, 0.008, 0.008, 0.024, 0.02, 0.0079, 0.004 => 0.0060

Dominant: 0, 0.004, 0, 0.004, 0.004, 0.004, 0.004, 0.004 => 0.003

Recessive: 0, 0, 0, 0.004, 0.004, 0.004, 0, 0.004 => 0.002

XOR: 0.031, 0.027, 0.024, 0.024 => 0.0265

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): 2, 3, 0, 0, 8.5, 251, 0, 0, 1, 1, 1, 1, 251, 7.5, 2.5, 3.5, 166.5, 249.5, 378.25, 441.25

Dominant Interaction Average True Positive Positions: 0, 1, 0, 341, 1, 1, 254, 476

Recessive Interaction Average True Positive Positions: 0, 0, 0, 251, 70, 131, 0, 445.5

XOR Interaction Average True Positive Positions: 63.75, 169.125, 137.5, 189

MDR proved versatile in being able to detect all interaction types.

Pairs Detected (# of SNP pairs correctly detected / total # of interacting disease SNP pairs) – only count if at least one pair was detected: 1/1, 1/1, 2/2, ½, 1/1, 1/1, 1/1, 1/1, ½, 2/2, 2/2, 2/2, 6/8, 5/8, 2/8, 1/8, 1/1, ½, 1/1, 1/1, ½, ½, ½, 1/1, 1/1, ½, 8/8, 7/8, 6/8, 6/8

The number of detected pairs was often close to the number of all interacting pairs present in the dataset.

Now let us examine when ­­­­MDR epistasis could detect multiplicative interaction.

Multiplicative Interaction Average True Positive Position Clustered by Interaction Alpha:

Interaction Alpha 1.25: 0, 0, 1, 3.5

Interaction Alpha 1.5: 0, 0, 1, 2.5

Interaction Alpha 2.0: 2, 8.5, 1, 251, 249.5, 441.25

Interaction Alpha 3.0: 3, 251, 1, 7.5, 166.5, 378.25

Clear trend here that interactions with larger effect were more easily detectable (for multiplicative interactions).

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, 1, 1, 251, 7.5, 2.5, 3.5, 166.5, 249.5

Impure: 2, 3, 0, 0, 8.5, 251, 0, 0, 378.25, 441.25

Pure epistasis is easier to detect as expected (for multiplicative interactions).