# Candidate eQTL detection for risk refinement using paired DNA-RNA panel data

Esther Hsiao[1], Linda Polfus[1], Jonathan Watterson[1], Hoda Mirsafian[1], David Burks[1], Adam Chamberlin[1], Matthew Schultz[1], Tina Pesaran[1], Donavan Cheng[1], Bojan Losic[1]
[1]Ambry Genetics, Inc.

## Abstract

Regulatory haplotypes modulate the expressivity and penetrance of hereditary cancer-risk mutations. We demonstrate expression quantitative trait loci (eQTL) detection using Ambry CancerNext and RNAInsight data from 57,898 patients, one of the largest clinical blood paired DNA-RNA datasets, validated against GTEx v8 whole blood eQTL data. After batch renormalization, 66 eQTLs were identified, showing strong additive effects ($R^2$ = 0.99, p <~ 1e-16). Select eQTLs correlated with age of onset in patients with loss-of-function mutations and variants of uncertain significance, even after adjusting for confounders. These findings support eQTL detection in clinical datasets and suggest regulatory haplotypes may refine risk assessment in hereditary cancer.

## Introduction

Inference of regulatory haplotypes which modify expressivity and penetrance of coding hereditary cancer-risk mutations is a key goal of precision medicine, potentially enabling deeper insight into individual risk profiles and therapeutic responses. Previous seminal work has already established that powerful functional readouts of latent regulatory variants acting on genes using phased haplotypes of coding variants and regulatory variants, including eQTL mapping in cis, can elucidate the enrichment of penetrance increasing configurations for pathogenic variants. In this work we demonstrate the feasibility of eQTL detection using Ambry CancerNext and RNAInsight data from 57,898 matched patients and present a preliminary eQTL analytic association analysis for potential risk refinement.

## Methods and Materials

- Ambry CancerNext and RNAInsight data from 57,898 samples were analyzed using genetic test results and matched gene expression measurements (TPM) respectively.
- GTEx v8 (whole blood) was used as a gold standard set to test for eQTL candidates in the Ambry data.
- eQTL concordance was defined as:
  1) significant expression change between wildtype and heterozygous alternative allele with Bonferroni adjusted p-value < 0.05, and
  2) having the same direction of effect changes between Ambry and GTEx data.
     - Effect size was calculated by batch-corrected TPM by accounting for sequencing run, gender, and ethnicity.
- Dosage ratio was calculated as $\frac{|WT-HOM|}{2\times|WT-HET|}$, where WT, HET, and HOM represent the median TPM of patients with wildtype, heterozygous, and homozygous alternative alleles, respectively. We defined "Additive" as ratio within 1 sd, "HOM-Dominant" as ratio > 1 + sd, and "WT-Dominant" as ratio < 1 – sd, where sd is the standard deviation (sd) of the 66 concordant eQTLs.
- Association and covariate renormalization analyses were performed using penalized linear models. Sequencing run, gender, and ethnicity were included as covariates. Standard statistical hypothesis testing methods were applied.

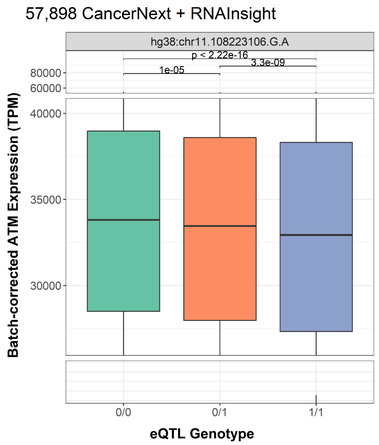


**Figure 1.** Example of a significant eQTL in ATM gene from Ambry CancerNext + RNAInsight data. The ATM batch-corrected expression decreased with each additional copy of the common SNP, which is consistent with the understanding that mutations in ATM can significantly increase the risk of developing breast cancer (BrCa).

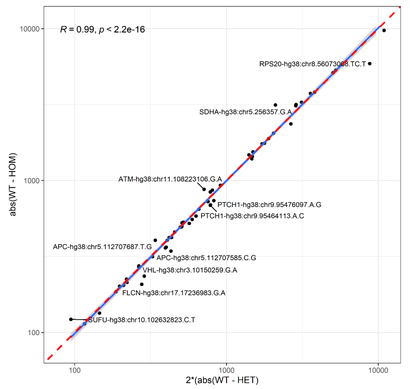


**Figure 2.** Dosage plot from 66 Ambry eQTLs that are concordant with GTEx whole blood data. The X-axis represents twice of the effect size between wildtype and heterozygous variant, and the Y-axis represents the effect size between wildtype and homozygous variant. Non-additive eQTLs are labeled by the gene names and hg38 genomic coordinates.

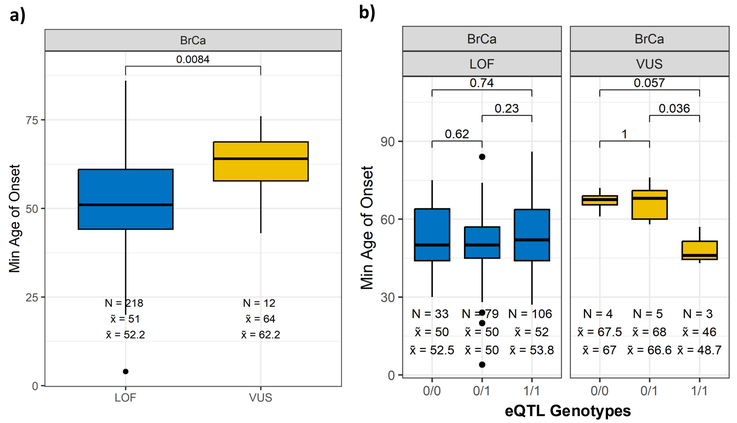


**Figure 3. a)** Age of onset comparison in Ambry BrCa patients from 57,898 CancerNext + RNAInsight data with VHL loss of function (LOF) variants or variants of unknown significance (VUS); **b)** same data as in a) but grouped the samples by the VHL eQTL genotypes (hg38:chr3.10150259.G>A). LOF: Samples with LOF variants in VHL and high-penetrance genes, such as BRCA1, BRCA2, TP53, PALB2, or CHEK2; VUS: Samples with any VUS variants in VHL.

## Results

- Ambry CancerNext+RNAInsight data from blood enabled eQTL detection. An example of significant eQTL in ATM is shown in **Figure 1**. The direction of expression change is consistent with our understanding that mutations in the ATM can significantly increase risk of developing breast cancer (BrCa).
- We observed a strongly additive dosage pattern ($R^2$ = 0.99, p <~ 1e-16) from 66 out of 96 eQTLs concordant with GTEx whole blood data after batch correction (**Figure 2**, **Table 1**). The 96 testable eQTLs are from 24 genes.
- Using select eQTLs we showed that age of onset in patients positive with loss of function mutations and variants of uncertain significance is correlated with eQTL genotype even after accounting for possible clinical and technical confounders (**Figure 3**).

**Table 1.** Ambry eQTLs concordant with GTEx blood eQTLs. Red: upregulating; blue: downregulating.

| Gene | Variant_hg38 | Group | Gene | Variant_hg38 | Group |
|---|---|---|---|---|---|
| SDHA | chr5.256357.G.A | HOM-Dominant | SDHA | chr5.235249.C.T | Additive |
| SUFU | chr10.102632823.C.T | HOM-Dominant | SMARCB1 | chr22.23834293.C.CG | Additive |
| ATM | chr11.108223106.G.A | HOM-Dominant | MSH2 | chr2.47385739.G.A | Additive |
| APC | chr5.112707687.T.G | HOM-Dominant | MSH2 | chr2.47373967.T.C | Additive |
| SDHA | chr5.224518.A.G | Additive | RAD51C | chr17.58358748.C.T | Additive |
| SDHA | chr5.226045.A.C | Additive | KIF1B | chr1.10361820.A.G | Additive |
| SDHA | chr5.228247.T.C | Additive | MSH2 | chr2.47385085.C.T | Additive |
| SDHA | chr5.251354.G.A | Additive | MSH3 | chr5.80654880.TGCAGCGGCTGCAGCGGCC.T | Additive |
| SDHA | chr5.251426.A.G | Additive | RAD51C | chr17.58357496.C.T | Additive |
| SDHA | chr5.254484.A.T | Additive | ATRIP | chr3.48446788.G.C | Additive |
| SDHA | chr5.233619.C.G | Additive | MSH3 | chr5.80654678.C.T | Additive |
| MSH6 | chr2.47790942.A.G | Additive | SDHC | chr1.161362556.C.G | Additive |
| MSH6 | chr2.47783419.C.A | Additive | RAD51C | chr17.58415439.T.C | Additive |
| MSH2 | chr2.47466649.A.T | Additive | MSH6 | chr2.47798625.C.T | Additive |
| CTNNA1 | chr5.138812254.A.C | Additive | BARD1 | chr2.214780821.C.G | Additive |
| SDHA | chr5.256394.G.A | Additive | MEN1 | chr11.64810148.G.C | Additive |
| BARD1 | chr2.214780740.C.G | Additive | KIF1B | chr1.10378629.C.T | Additive |
| MSH2 | chr2.47386731.T.C | Additive | SDHC | chr1.161314434.A.AGT | Additive |
| MSH3 | chr5.80854162.A.G | Additive | ATM | chr11.108270974.G.T | Additive |
| MLH3 | chr14.75017109.T.C | Additive | MLH1 | chr3.36993455.G.A | Additive |
| MLH3 | chr14.75047125.G.A | Additive | SUFU | chr10.102627262.T.C | Additive |
| RAD51C | chr17.58692618.C.T | Additive | LZTR1 | chr22.20983036.G.A | Additive |
| BARD1 | chr2.214809599.C.G | Additive | LZTR1 | chr22.20985799.T.C | Additive |
| BARD1 | chr2.214809617.A.G | Additive | MSH3 | chr5.80654689.C.T | Additive |
| CTNNA1 | chr5.138886300.G.C | Additive | PTCH1 | chr9.95464025.C.T | Additive |
| CTNNA1 | chr5.138930857.G.A | Additive | RPS20 | chr8.56074389.T.C | Additive |
| CTNNA1 | chr5.138934602.T.TATC | Additive | FLCN | chr17.17231866.G.A | Additive |
| CTNNA1 | chr5.138934853.T.A | Additive | PTCH1 | chr9.95476097.A.G | WT-Dominant |
| MSH3 | chr5.80654962.A.G | Additive | PTCH1 | chr9.95464113.A.C | WT-Dominant |
| PMS2 | chr7.5987357.G.A | Additive | VHL | chr3.10150259.G.A | WT-Dominant |
| BARD1 | chr2.214809647.C.T | Additive | APC | chr5.112707585.C.G | WT-Dominant |
| MSH2 | chr2.47385725.C.CTG | Additive | FLCN | chr17.17236983.G.A | WT-Dominant |
| SMARCB1 | chr22.23825326.G.A | Additive | RPS20 | chr8.56073068.TC.T | WT-Dominant |

## Summary & Conclusion

Our findings highlight the value of integrating regulatory variant analysis with clinical genetic testing to refine individual cancer risk profiles. The identification of eQTLs in a large clinical cohort underscores their potential to modify disease penetrance and expressivity, offering new avenues for precision medicine. These results support the use of eQTL mapping for risk refinement and suggest that germline eQTL variation may provide mechanistic prior information that complements tumor based HRD RNA signatures, enabling more refined identification of HRD-like biology beyond canonical BRCA1/2 alterations.

## Contact

Bojan Losic
Ambry Genetics , Aliso Viejo, CA, USA
Email: blosic@ambrygen.com

## References

1. Castel, S. E., et al. "Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk." Nature Genetics 50, no. 9 (2018): 1327-1334.
2. The GTEx Consortium. "The GTEx Consortium atlas of genetic regulatory effects across human tissues." Science 369, no. 6509 (2020): 1318-1330.
3. Li, Q., et al. "Expression quantitative trait loci (eQTLs) in cancer: implications for clinical genomics." Cancer Research 81, no. 16 (2021): 4375-4382.
4. Dorling, L., et al. "Breast cancer risk variants in the ATM gene: a case-control study and meta-analysis." Journal of the National Cancer Institute 114, no. 2 (2022): 256-265.