**Working Title:**

**Integrated omics analysis in Parkinson’s disease: genome-wide analysis of epigenetics and metabolomics provides links to Parkinson’ disease**

Kimberly C Paul\*, Gary Chen\*, Cynthia Kusters, Janet Sinsheimer, Jeff Bronstein, Beate Ritz

\*Co-First Author

**Objective:**

1. Identify blood-based methylation and metabolomic quantitative trait loci (meQTLs and metQTLs) and integrate them with PD-related genetic variants, epigenetic variants, and metabolites discovered in GWAS, EWAS, and MWAS.
2. Assess enrichment of QTLs observed in PD patients: (a) did PD patients have more QTLs than expected; (b) are QTLs enriched for known biologic pathways (KEGG/GO/etc GSEA)

“Identifying meQTL variants and linking them to disease-associated genetic variants from GWAS would pinpoint molecular mechanisms underlying genetic susceptibility to human diseases that are due, at least in part, to altered epigenetic regulation. Additionally, it could help explain the molecular consequences of non-protein-coding, disease-associated genetic variants from GWAS.”

<https://www.nature.com/articles/s41467-019-12228-z>

**Analysis Outline:**

1. Perform genome-wide association testing of genetic variants with whole blood DNA methylation and untargeted serum-based metabolites from 750 participants of the PEG Parkinson’s Study (541 discovery cohort, 209 replication cohort) to identify cis-meQTLs (methylation QTLs), trans-meQTLs, and metQTLs (metabolite QTLs).
   1. QTL hotspots (regions of the genome that have more QTLs than expected/others)
   2. Colocalization of QTLs (meQTLs, metQTLs, and eQTLs from GTeX), what proportion of QTLs influence a single omic layer versus influencing multiple (ie most meQTLs are also eQTLs, metQTLs?)
2. Link the QTLs to GWAS summary statistics for Parkinson’s disease from the meta-GWAS
   1. Assess whether there are more QTLs among GWAS SNPs than expected (ie x% of all SNPs are QTLs, y% of GWAS SNPs (p<0.05) are QTLs, overrepresented?)
   2. Do PD genes have more QTLs among patients than expected
   3. Colocalization? which GWAS SNPs colocalize with a QTL
   4. Mendelian randomization (MR) to identify putatively causal CpGs for PD? (https://www.nature.com/articles/s41467-019-12228-z)
3. Enrichment of QTLs for pathways, etc (KEGG/GO/etc): meQTLs only, metQTLs only, and combined?

PPMI as replication cohort now? Or leave out as they don’t have the metabolomics?

Results so far showing a lot of immune pathways, should we add a small section of how QTLs are related to immunoplex panel (measured levels of 38 immune factors: cytokines, chemokines, growth factors, etc.)?

**Results:**

(Sup?) Table 1: Study Information N (age, sex, AIMs ancestry (though we are limiting to Euro, still can show fractional ancestry), etc)

Table 2: meQTL output (large table of all SNP~CpG results, FDR<0.05 or maybe even p<0.05 for supplement)

PEG1 discovery, PEG2 test

Include indicator for cis or trans

Table 3: meQTL annotations, including PD GWAS SNP/p-value if so, gene meQTL located in, and gene region (TSS, UTR, etc)

Table 4: meQTL enrichment, including PD GWAS SNPs, gene / genome region, PD pathways, GO terms (GREAT analysis)

Table 5: known meQTL?

Link to BIOS meQTLs

known eQTL (GTex)

do we have any unique meQTLs?

**Table 1 Study Information**

|  |  |  |
| --- | --- | --- |
|  | **PEG 1** | **PEG 2** |
| Sample size (female/male) | 231/310 | 64/145 |
| Age (mean, range) | 69 (35-92) | 71 (46-92) |

**Supplementary Table 1** Fractional Ancestry Information among four clusters1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **PEG 1 clusters** | | | | **PEG 2 clusters** | | | |
| Self-reported ethnicity | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Latino | 0.038 | 0.490 | 0.470 | 0.002 | 0.034 | 0.492 | 0.472 | 0.001 |
| White | 0.011 | 0.007 | 0.976 | 0.006 | 0.009 | 0.017 | 0.970 | 0.004 |
| Asian | 0.280 | 0.717 | 0.002 | 0.001 | 0.279 | 0.718 | 0.002 | 0.001 |
| Native American | 0.024 | 0.008 | 0.965 | 0.003 | 0.005 | 0.003 | 0.985 | 0.008 |
| Other |  |  |  |  | 0.003 | 0.002 | 0.991 | 0.004 |

1 ethnicity clusters defined in STRUCTURE software program

**Table 2: List of ten meQTLs ranked by statistical evidence, discovered in PEG1 and replicated in PEG2**

(Supplementary table lists all results at FDR<0.05)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **type** | **snp** | **allele** | **gene** | **cohort** | **p-value** | **FDR** | **beta** | **cohort** | **p-value** | **FDR** | **beta** |
| cis | GSA-rs1040961 | G | cg17707870 | peg1 | 2.65E-242 | 2.50E-234 | 0.463637 | peg2 | 1.25E-83 | 1.23E-76 | 0.48246 |
| cis | rs10010994 | C | cg17858192 | peg1 | 9.19E-232 | 4.35E-224 | 0.384347 | peg2 | 9.10E-75 | 2.70E-68 | 0.38966 |
| cis | exm2267473 | G | cg09084244 | peg1 | 1.33E-217 | 4.20E-210 | 0.425392 | peg2 | 2.73E-72 | 5.70E-66 | 0.415686 |
| cis | rs10184015 | A | cg02502145 | peg1 | 6.88E-215 | 1.63E-207 | 0.458718 | peg2 | 9.63E-80 | 4.72E-73 | 0.483753 |
| cis | rs2532925 | G | cg04145681 | peg1 | 9.34E-215 | 1.77E-207 | 0.473452 | peg2 | 2.21E-76 | 7.74E-70 | 0.458496 |
| cis | GSA-rs1035142 | T | cg07227024 | peg1 | 1.79E-205 | 2.83E-198 | 0.429669 | peg2 | 1.97E-75 | 6.23E-69 | 0.438633 |
| cis | GSA-rs10750097 | G | cg12556569 | peg1 | 1.30E-203 | 1.75E-196 | 0.384724 | peg2 | 3.05E-80 | 1.57E-73 | 0.408519 |
| cis | rs2883456 | C | cg11144103 | peg1 | 2.55E-203 | 3.01E-196 | 0.437497 | peg2 | 1.28E-93 | 4.18E-86 | 0.46732 |
| cis | rs1939015 | G | cg10306192 | peg1 | 7.31E-200 | 7.68E-193 | 0.493061 | peg2 | 3.28E-93 | 8.04E-86 | 0.539282 |
| cis | rs1043793 | A | cg02502145 | peg1 | 2.07E-198 | 1.95E-191 | 0.452547 | peg2 | 1.71E-47 | 3.06E-42 | 0.439154 |

**Table 3: List of ten meQTLs ranked by statistical evidence, discovered in PEG1 and replicated by Nall GWAS**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **meQTL evidence** | | | **GWAS evidence** | | | |
| **type** | **snp** | **allele** | **probe** | **cohort** | **pvalue** | **FDR** | **beta** | **beta** | **se** | **pvalue** | **gene** |
| cis | rs72660967 | T | cg06961873 | peg1 | 1.79E-158 | 2.92E-152 | 0.408418 | 0.0439 | 0.019 | 0.0206 | missing |
| cis | rs10903129 | A | cg06961873 | peg1 | 7.88E-158 | 1.22E-151 | -0.409488 | -0.0435 | 0.0168 | 0.009549 | missing |
| cis | rs3087801 | G | cg10776061 | peg1 | 3.55E-157 | 5.42E-151 | 0.37696 | -0.0401 | 0.0185 | 0.03017 | missing |
| cis | rs9276490 | A | cg18572898 | peg1 | 1.70E-153 | 1.78E-147 | 0.41004 | -0.0448 | 0.0224 | 0.04595 | missing |
| cis | rs9649865 | A | cg25543264 | peg1 | 4.91E-152 | 4.65E-146 | 0.385467 | 0.0505 | 0.0223 | 0.02349 | missing |
| cis | rs166849 | G | cg13143872 | peg1 | 5.40E-150 | 4.65E-144 | 0.312391 | -0.0391 | 0.0192 | 0.04131 | missing |
| cis | rs10191694 | A | cg15147113 | peg1 | 2.74E-148 | 2.23E-142 | 0.108973 | 0.0717 | 0.0235 | 0.002248 | missing |
| cis | rs7799245 | G | cg11957130 | peg1 | 7.39E-146 | 4.95E-140 | 0.440158 | 0.0685 | 0.0246 | 0.005368 | missing |
| cis | rs9276490 | A | cg07389699 | peg1 | 1.04E-145 | 6.40E-140 | 0.342294 | -0.0448 | 0.0224 | 0.04595 | missing |
| cis | rs9689 | A | cg06961873 | peg1 | 2.71E-135 | 1.34E-129 | -0.392911 | -0.0392 | 0.0175 | 0.02491 | missing |

**Table 4: trans meQTL hotspots**

|  |  |  |
| --- | --- | --- |
| **SNP ID** | **Associated probes in PEG1** | **Associated probes in PEG2** |
| rs4870921\_T | 1 | 11497 |
| rs1430849\_T | 1 | 4215 |
| rs4058295\_C | 1 | 637 |
| GSA-rs115731818\_G | 2 | 510 |
| rs12657428\_T | 1 | 350 |
| rs3803354\_C | 346 | 1 |
| rs1530942\_A | 335 | 1 |
| rs6446023\_A | 7 | 145 |
| rs864072\_A | 1 | 123 |
| rs2603158\_G | 111 | 9 |
| rs341081\_T | 1 | 117 |
| rs79067366\_C | 1 | 110 |
| rs4774314\_C | 38 | 67 |
| rs6808554\_C | 93 | 1 |
| rs7809957\_T | 1 | 91 |
| rs12657183\_G | 1 | 91 |
| rs17079652\_A | 1 | 89 |
| GSA-rs5764394\_A | 79 | 1 |
| rs1978619\_T | 71 | 4 |
| rs73156693\_C | 1 | 70 |

**Table 5 meQTLs that are also metQTLs in PEG1**

|  |  |  |
| --- | --- | --- |
| snp\_id | allele | associated\_probes |
| rs4246215 | G | 30 |
| rs174550 | T | 25 |
| rs174537 | G | 25 |
| rs174536 | A | 25 |
| rs174528 | T | 25 |
| rs174546 | C | 25 |
| rs174547 | T | 25 |
| rs174576 | C | 25 |
| rs1535 | A | 20 |
| rs174583 | C | 20 |
| rs174529 | T | 20 |
| rs174535 | T | 20 |
| rs174538 | G | 20 |
| rs174574 | A | 15 |
| rs102275 | T | 15 |
| rs108499 | C | 15 |
| rs174577 | C | 15 |
| rs174534 | A | 12 |
| rs174601 | C | 12 |
| rs2147896 | A | 10 |
| rs4345897 | A | 10 |
| rs4539242 | T | 10 |
| GSA-rs11189581 | T | 10 |
| rs7072216 | T | 10 |
| GSA-rs174600 | T | 8 |
| GSA-rs17201602 | C | 7 |
| rs34878747 | C | 7 |
| rs7915108 | T | 6 |
| rs174549 | G | 5 |
| rs174555 | T | 5 |
| rs174556 | C | 5 |
| rs495828 | G | 3 |
| rs579459 | T | 3 |
| rs635634 | C | 3 |
| rs651007 | C | 3 |
| 9:136146597-C-T | C | 3 |
| rs2066938 | A | 3 |
| rs2014355 | T | 3 |
| rs17109597 | G | 3 |
| rs10883094 | G | 3 |
| exm-rs507666 | G | 3 |
| 9:136145425-C-A | C | 3 |
| rs1951033 | T | 2 |
| rs10138605 | C | 2 |
| rs4320932 | T | 1 |
| rs12682122 | A | 1 |
| rs632111 | A | 1 |
| exm1487912 | G | 1 |
| rs8050881 | A | 1 |

**Table 6 meQTLs that are also metQTLs in PEG2**

|  |  |  |
| --- | --- | --- |
| snp\_id | allele | associated\_probes |
| rs58314476 | A | 16 |
| GSA-rs4734295 | A | 2 |
| rs7743761 | C | 2 |
| rs1240707 | C | 1 |
| rs2075516 | T | 1 |
| rs4712969 | A | 1 |
| rs9820731 | A | 1 |

**Figure 1: meQTL enrichment in GO Biological Process (GREAT analysis)**

Chart, diagram

Description automatically generated

**Table 4: List of ten meQTLs ranked by statistical evidence, also found in BIOS repository**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **SNP** | | **Probe** | |  |  |  |  |  |
| **Chr** | **ID** | **Position** | **ID** | **Position** | **Alleles** | **Tested Allele** | **Z-score** | **p-value** | **HGNCName** |
| 13 | rs11616903 | 50424677 | cg04520693 | 50425377 | C/T | T | 72.6668802 | 3.27167E-310 | RNY4P30 |
| 4 | rs4406078 | 57521325 | cg19978674 | 57523802 | T/C | T | 72.1343579 | 3.27167E-310 | HOPX |
| 2 | rs12475055 | 102878891 | cg23719516 | 102870432 | A/C | A | -71.8785276 | 3.27167E-310 | AC007248.6 |
| 11 | rs515449 | 122047290 | cg13160852 | 122047052 | G/A | A | -71.3891567 | 3.27167E-310 | RP11-820L6.1 |
| 8 | rs10086636 | 126604637 | cg08775595 | 126608666 | C/T | T | -71.3193873 | 3.27167E-310 | RP11-136O12.2 |
| 4 | rs11730732 | 119761678 | cg11663691 | 119759924 | A/G | A | 71.2074002 | 3.27167E-310 | SEC24D |
| 8 | rs10090179 | 144339053 | cg19504605 | 144339082 | G/A | A | -71.1133419 | 3.27167E-310 | ZFP41 |
| 8 | rs1454616 | 126611214 | cg08775595 | 126608666 | T/C | C | -71.0101274 | 3.27167E-310 | RP11-136O12.2 |
| 14 | rs10483261 | 22356551 | cg16702660 | 22361306 | T/C | C | -70.9943176 | 3.27167E-310 | TRAV8-4 |
| 17 | rs9894429 | 79596811 | cg21028142 | 79581735 | C/T | T | 70.9090052 | 3.27167E-310 | NPLOC4 |