Running Metasoft on all significant single tissue SNP-gene pair eQTLs using GTEx release v6 data

Xiao Li, Hailei Zhang, Ayellet Segrè

GTEx LDACC
Broad Institute

April 28, 2016

Metasoft

We used the new random effects model in Metasoft (called RE2), designed to find loci with effects that may have heterogeneity between datasets/tissues (assumes estimates are independent)
[Buhm Han and Eleazar Eskin, "Random-Effects Model Aimed at Discovering Associations in Meta-Analysis of Genome-wide Association Studies", The American Journal of Human Genetics (2011) 88, 586-598.]
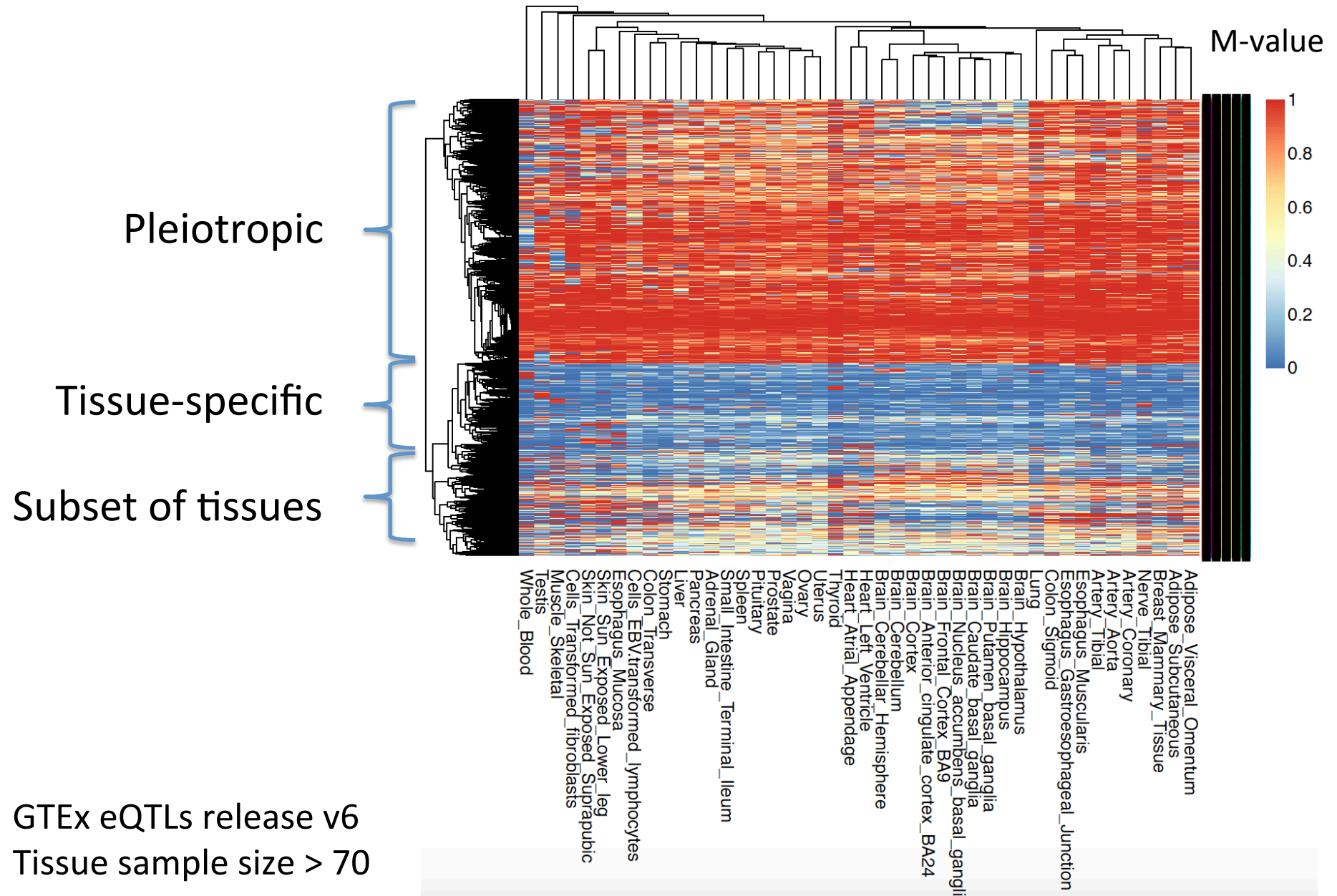
Metasoft computes an M-value, a metric that measures probability that the effects will exist or not in a given dataset.
[Buhm Han and Eleazar Eskin, "Interpreting Meta-Analysis of Genome-wide Association Studies", PLoS Genetics (2012)]

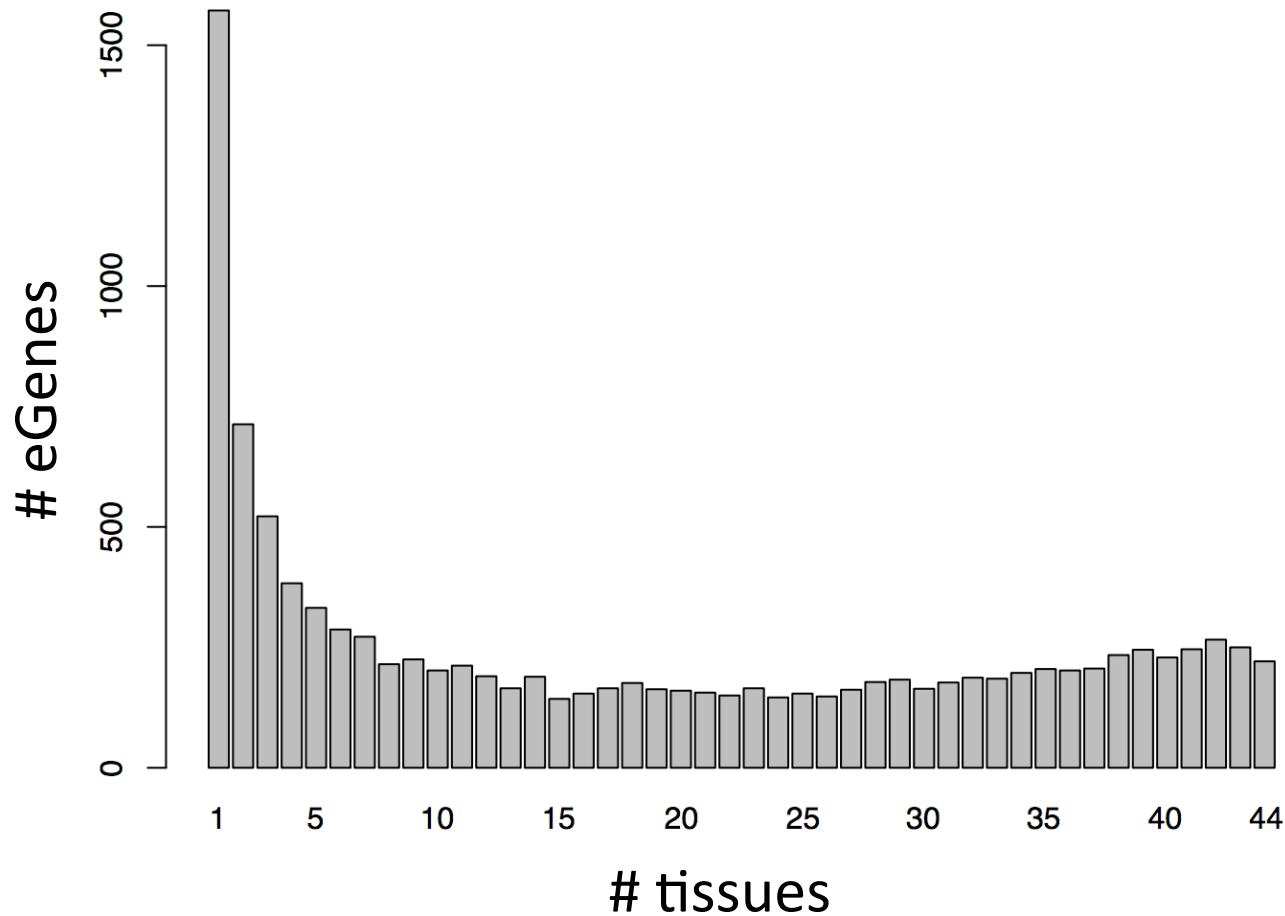# Running Metasoft on all significant single tissue SNP-gene pair eQTLs using GTEx release v6 data

- We selected SNP gene pairs that are significant in at least one of 44 tissues at FDR<5% based on Matrix eQTL analysis as the input for Metasoft.

- We calculated M-values for each significant SNP-gene pairs.

- For each gene, we selected the best SNP for that gene by choosing the one with smallest Pval_RE2.

Clustering of best eQTL per eGene across 44 GTEx tissues

Pleiotropic

Tissue-specific

Subset of tissues

M-value

GTEx eQTLs release v6
Tissue sample size > 70

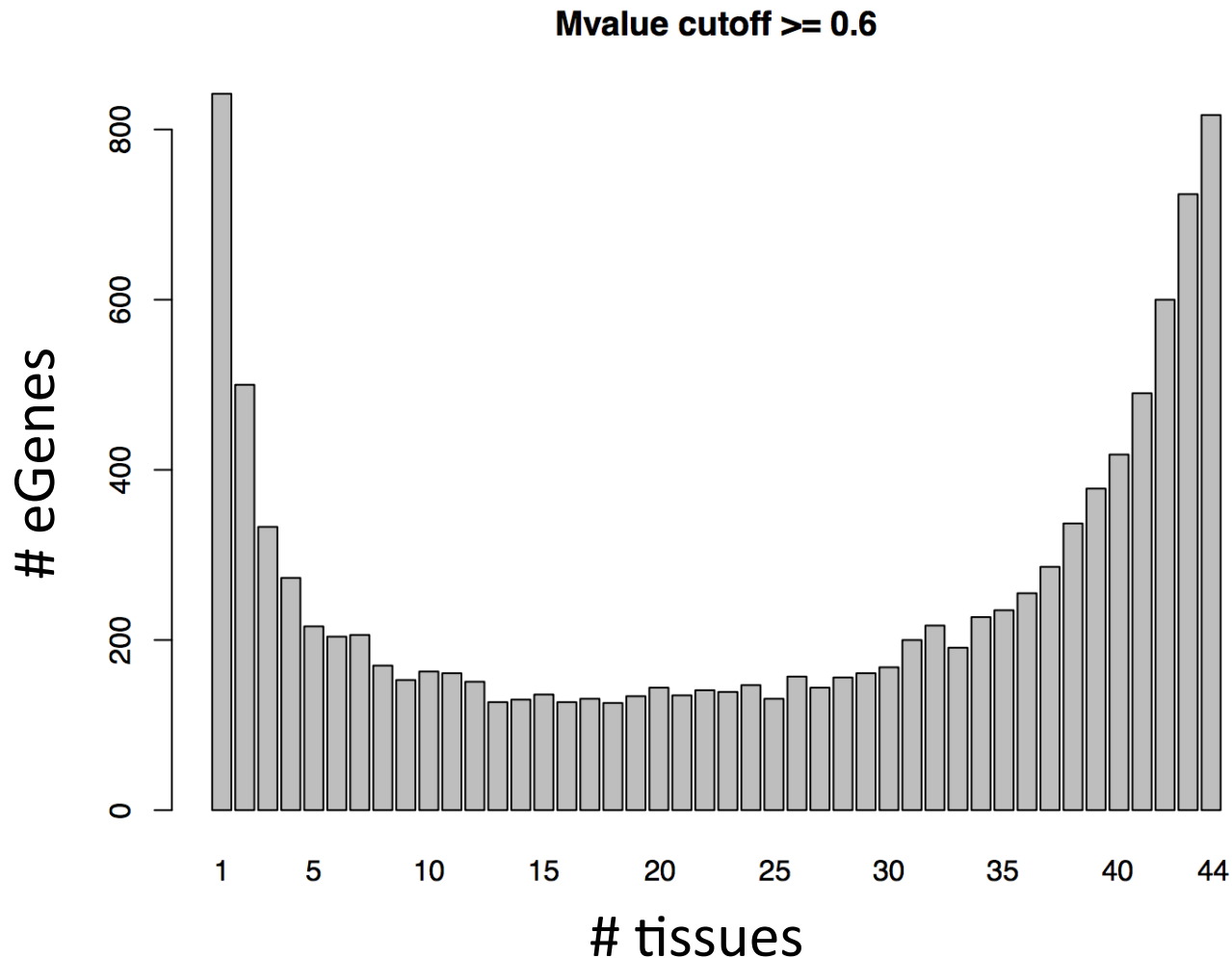# Best eQTL per eGene sharing across 44 GTEx tissues (conservative view)

Metasoft (Buhm Han, Eleaazr Eskin, UCLA)
Xiao Li , Broad Institute

eQTL sharing is likely to increase with increased sample sizes

Mvalue cutoff >= 0.6

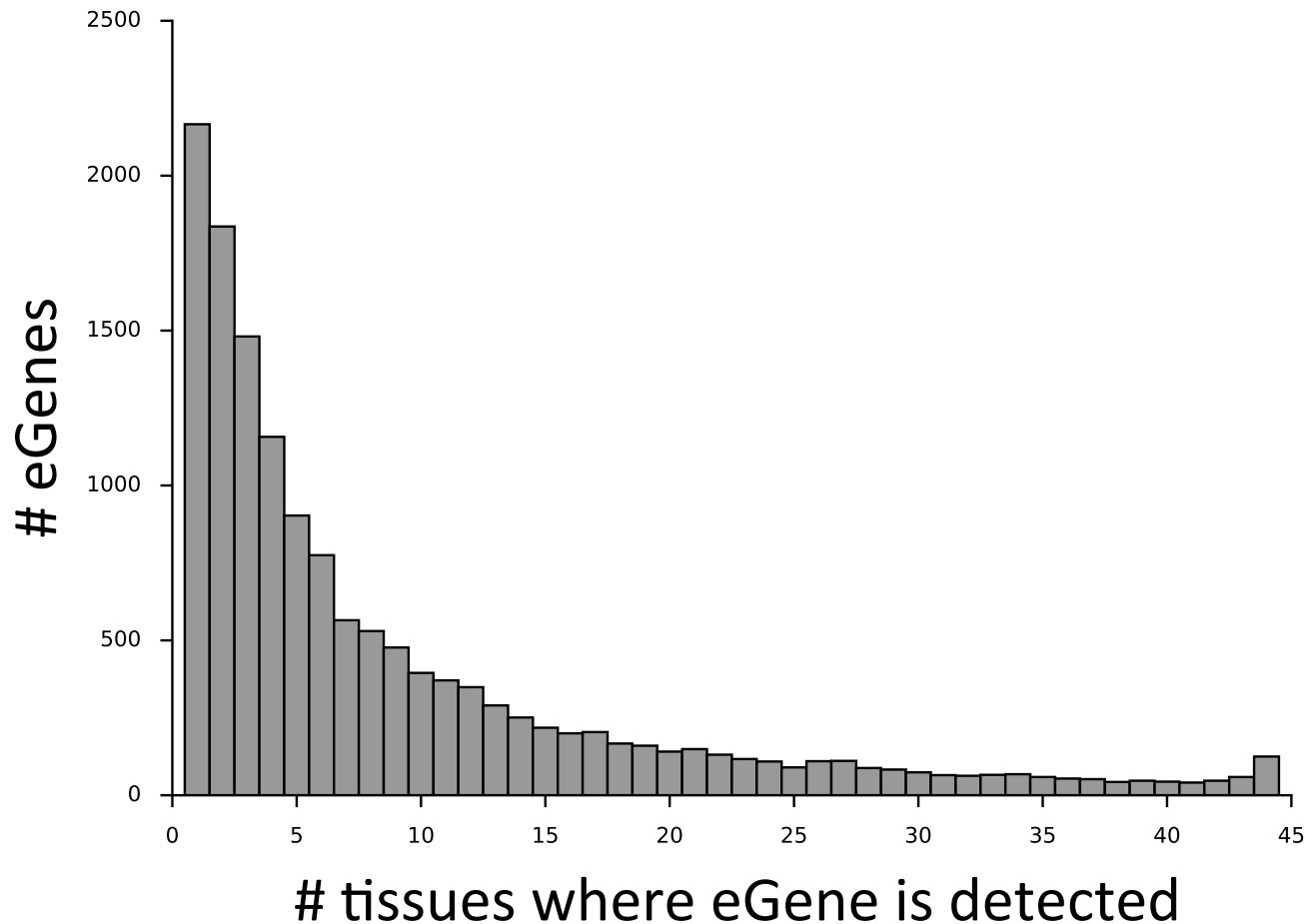**Metasoft** (Buhm Han, Eleaazr Eskin, UCLA)
Xiao Li , Broad Institute

eGene sharing across 44 GTEx tissues
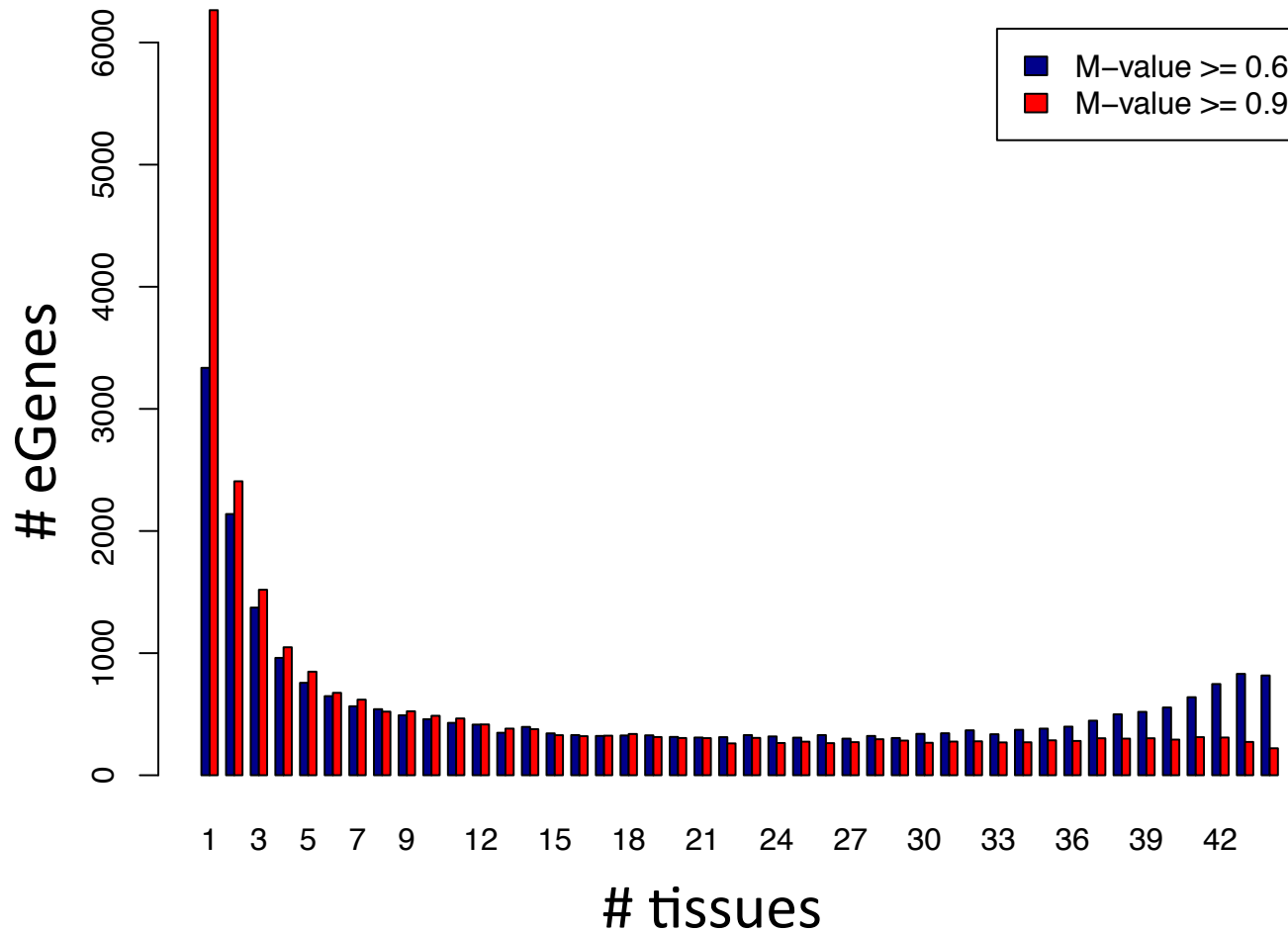
Metasoft (Buhm Han, Eleaazr Eskin, UCLA)
Xiao Li , Broad Institute

eGene sharing across 44 GTEx tissues using Matrix eQTL results (FDR<5%)

Comparison of multi-tissue Metasoft results to single-tissue Matrix eQTL results

Comparison of Metasoft to Matrix eQTL results

M-value ≥ 0.9

tissure.type	#eGene.Matrix.eQTL	#eGene.Metasoft	#overlap.eGene	%overlap.eGene/MatrixeQTL	%overlap.eGene/Metasoft	#eGene.unique.MatrixeQTL	#eGene.unique.Metasoft	#overlap.eQTL	#eQTL.unique.Metasoft	#eQTL.unique.MatrixeQTL	%overlap.eQTL/Metasoft	%overlap.eQTL/MatrixeQTL
Adipose_Subcutaneous	8500	10165	6954	81.81	68.41	1546	3211	686	9479	7814	6.75	8.07
Adipose_Visceral_Omentum	4301	8239	3657	85.03	44.39	644	4582	265	7974	4036	3.22	6.16
Adrenal_Gland	3259	6917	2553	78.34	36.91	706	4364	154	6763	3105	2.23	4.73
Artery_Aorta	6220	8836	5016	80.64	56.77	1204	3820	412	8424	5808	4.66	6.62
Artery_Coronary	2363	6952	2030	85.91	29.2	333	4922	109	6843	2254	1.57	4.61
Artery_Tibial	8056	9414	6466	80.26	68.68	1590	2948	583	8831	7473	6.19	7.24
Brain_Anterior_cingulate_cortex_BA24	1212	5088	1010	83.33	19.85	202	4078	43	5045	1169	0.85	3.55
Brain_Caudate_basal_ganglia	2447	6285	1966	80.34	31.28	481	4319	119	6166	2328	1.89	4.86
Brain_Cerebellar_Hemisphere	3251	6056	2417	74.35	39.91	834	3639	180	5876	3071	2.97	5.54
Brain_Cerebellum	4163	6836	3112	74.75	45.52	1051	3724	217	6619	3946	3.17	5.21
Brain_Cortex	2567	6181	2062	80.33	33.36	505	4119	124	6057	2443	2.01	4.83
Brain_Frontal_Cortex_BA9	2009	5719	1621	80.69	28.34	388	4098	81	5638	1928	1.42	4.03
Brain_Hippocampus	1134	5106	939	82.8	18.39	195	4167	33	5073	1101	0.65	2.91
Brain_Hypothalamus	1157	5187	993	85.83	19.14	164	4194	39	5148	1118	0.75	3.37
Brain_Nucleus_accumbens_basal_ganglia	2019	5954	1655	81.97	27.8	364	4299	88	5866	1931	1.48	4.36
Brain_Putamen_basal_ganglia	1588	5561	1309	82.43	23.54	279	4252	68	5493	1520	1.22	4.28
Breast_Mammary_Tissue	4140	8375	3593	86.79	42.9	547	4782	234	8141	3906	2.79	5.65
Cells_EBV-transformed_lymphocytes	2957	5461	2189	74.03	40.08	768	3272	153	5308	2804	2.8	5.17
Cells_Transformed_fibroblasts	8760	8043	6094	69.57	75.77	2666	1949	633	7410	8127	7.87	7.23
Colon_Sigmoid	2882	7209	2505	86.92	34.75	377	4704	133	7076	2749	1.84	4.61
Colon_Transverse	4446	7801	3646	82.01	46.74	800	4155	311	7490	4135	3.99	7
Esophagus_Gastroesophageal_Junction	2751	7431	2419	87.93	32.55	332	5012	123	7308	2628	1.66	4.47
Esophagus_Mucosa	7416	8929	5670	76.46	63.5	1746	3259	476	8453	6940	5.33	6.42
Esophagus_Muscularis	6916	9549	5948	86	62.29	968	3601	627	8922	6289	6.57	9.07
Heart_Atrial_Appendage	3929	7500	3236	82.36	43.15	693	4264	237	7263	3692	3.16	6.03
Heart_Left_Ventricle	4417	7288	3515	79.58	48.23	902	3773	234	7054	4183	3.21	5.3
Liver	1628	5051	1194	73.34	23.64	434	3857	53	4998	1575	1.05	3.26
Lung	7236	9461	5869	81.11	62.03	1367	3592	557	8904	6679	5.89	7.7
Muscle_Skeletal	7082	7911	5249	74.12	66.35	1833	2662	499	7412	6583	6.31	7.05
Nerve_Tibial	9860	10875	7964	80.77	73.23	1896	2911	850	10025	9010	7.82	8.62
Ovary	1583	5859	1343	84.84	22.92	240	4516	66	5793	1517	1.13	4.17
Pancreas	4301	7311	3288	76.45	44.97	1013	4023	206	7105	4095	2.82	4.79
Pituitary	2160	6321	1808	83.7	28.6	352	4513	96	6225	2064	1.52	4.44
Prostate	1462	6102	1249	85.43	20.47	213	4853	47	6055	1415	0.77	3.21
Skin_Not_Sun_Exposed_Suprapubic	5491	8742	4557	82.99	52.13	934	4185	354	8388	5137	4.05	6.45
Skin_Sun_Exposed_Lower_leg	8567	10240	7147	83.42	69.79	1420	3093	820	9420	7747	8.01	9.57
Small_Intestine_Terminal_Ileum	1356	5730	1138	83.92	19.86	218	4592	49	5681	1307	0.86	3.61
Spleen	2754	6599	2168	78.72	32.85	586	4431	131	6468	2623	1.99	4.76
Stomach	3438	7394	2880	83.77	38.95	558	4514	212	7182	3226	2.87	6.17
Testis	9009	8709	5719	63.48	65.67	3290	2990	661	8048	8348	7.59	7.34
Thyroid	9937	10870	7993	80.44	73.53	1944	2877	895	9975	9042	8.23	9.01
Uterus	917	5402	777	84.73	14.38	140	4625	20	5382	897	0.37	2.18
Vagina	840	5294	734	87.38	13.86	106	4560	25	5269	815	0.47	2.98
Whole_Blood	6784	6264	4315	63.61	68.89	2469	1949	434	5830	6350	6.93	6.4
Average	**4256**	**7277.7**	**3362.9**	**80.5**	**42.6**	**893.1**	**3914.8**	**280.4**	**6997.3**	**3975.6**	**3.4**	**5.5**

Hailei Zhang, Broad Institute

# Comparison of Metasoft to Matrix eQTL results

Average across 44 tissues	Metasoft M-value ≥ 0.9	Metasoft M-value ≥ 0.6
# eGenes, Matrix eQTL (FDR<5%)	4,256	4,256
# eGenes, Metasoft	7,278	10,413
# overlapping eGenes	3,363	3,591
% overlapping eGenes, of MatrixeQTL	80.5%	86%
% overlapping eGenes, of Metasoft	42.6%	33%
# (%) eGenes unique to MatrixeQTL	893 (20.1%)	665 (15.6%)
# (%) eGenes unique to Metasoft	3,915 (54%)	6,822 (65.5%)
# overlapping SNP-gene pairs	280	290
% overlapping SNP-gene pairs, of Metasoft	3.4%	2.6%
% overlapping SNP-gene pairs, of MatrixeQTL	5.5%	5.7%

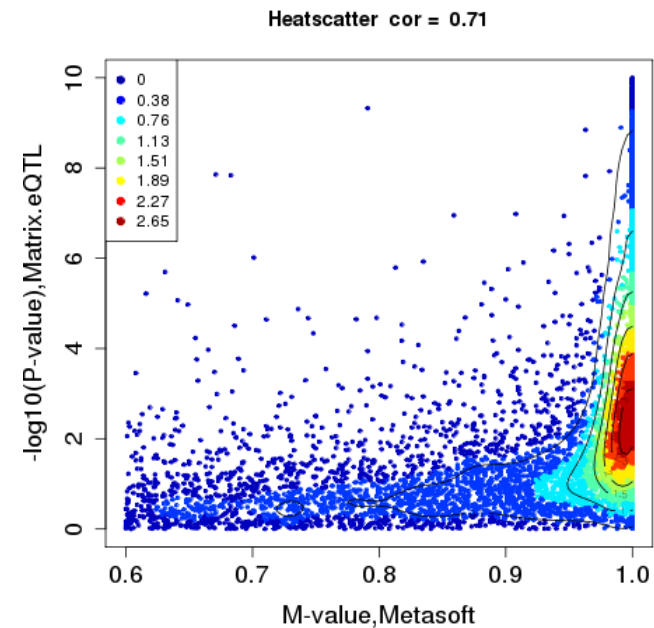Hailei Zhang, Broad Institute

Comparison of best eQTL per eGene between Metasoft and Matrix eQTL
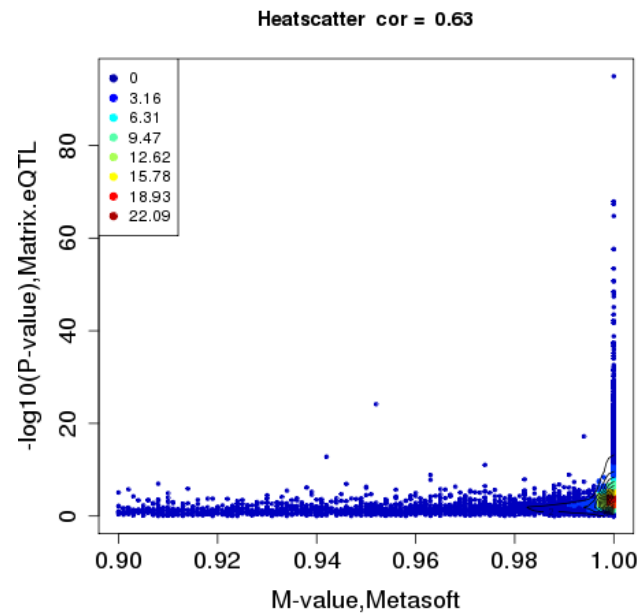
M-value ≥ 0.9
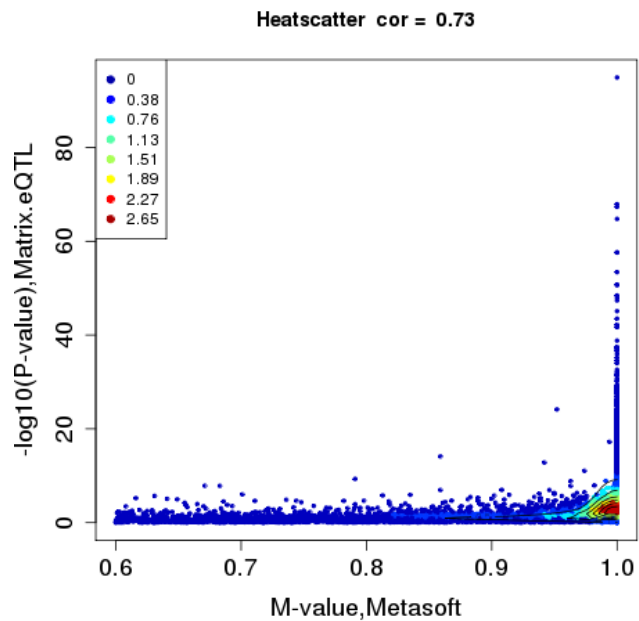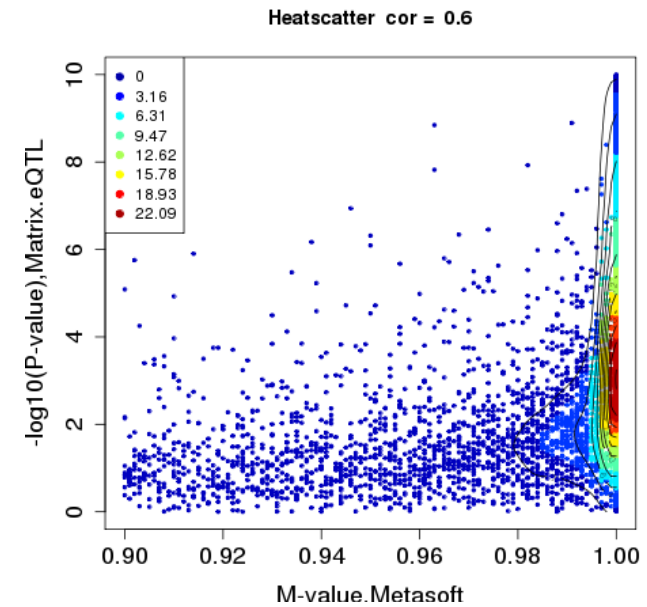
M-value ≥ 0.6

Comparison of best eQTL per eGene between Metasoft (M-value≥0.9) and Matrix eQTL

Tissue	Pi1	Spearman's correlation coefficent
Adipose_Subcutaneous	0.977	0.675
Adipose_Visceral_Omentum	0.973	0.701
Adrenal_Gland	0.973	0.694
Artery_Aorta	0.987	0.71
Artery_Coronary	0.957	0.676
Artery_Tibial	0.993	0.685
Brain_Anterior_cingulate_cortex_BA24	0.931	0.644
Brain_Caudate_basal_ganglia	0.954	0.691
Brain_Cerebellar_Hemisphere	0.958	0.692
Brain_Cerebellum	0.972	0.695
Brain_Cortex	0.954	0.694
Brain_Frontal_Cortex_BA9	0.94	0.663
Brain_Hippocampus	0.921	0.65
Brain_Hypothalamus	0.919	0.644
Brain_Nucleus_accumbens_basal_ganglia	0.948	0.687
Brain_Putamen_basal_ganglia	0.948	0.668
Breast_Mammary_Tissue	0.987	0.702
Cells_EBV-transformed_lymphocytes	0.977	0.702
Cells_Transformed_fibroblasts	0.988	0.648
Colon_Sigmoid	0.972	0.721
Colon_Transverse	0.967	0.693
Esophagus_Gastroesophageal_Junction	0.968	0.7
Esophagus_Mucosa	0.989	0.693
Esophagus_Muscularis	0.99	0.705
Heart_Atrial_Appendage	0.977	0.7
Heart_Left_Ventricle	0.974	0.68
Liver	0.946	0.66
Lung	0.986	0.692
Muscle_Skeletal	0.99	0.667
Nerve_Tibial	0.992	0.691
Ovary	0.94	0.675
Pancreas	0.976	0.713
Pituitary	0.948	0.692
Prostate	0.942	0.655
Skin_Not_Sun_Exposed_Suprapubic	0.987	0.707
Skin_Sun_Exposed_Lower_leg	0.998	0.68
Small_Intestine_Terminal_Ileum	0.915	0.631
Spleen	0.975	0.695
Stomach	0.968	0.702
Testis	0.991	0.672
Thyroid	0.986	0.662
Uterus	0.933	0.632
Vagina	0.941	0.633
Whole_Blood	0.976	0.627
Mean:	**0.966**	**0.68**
Median:	**0.966**	**0.68**

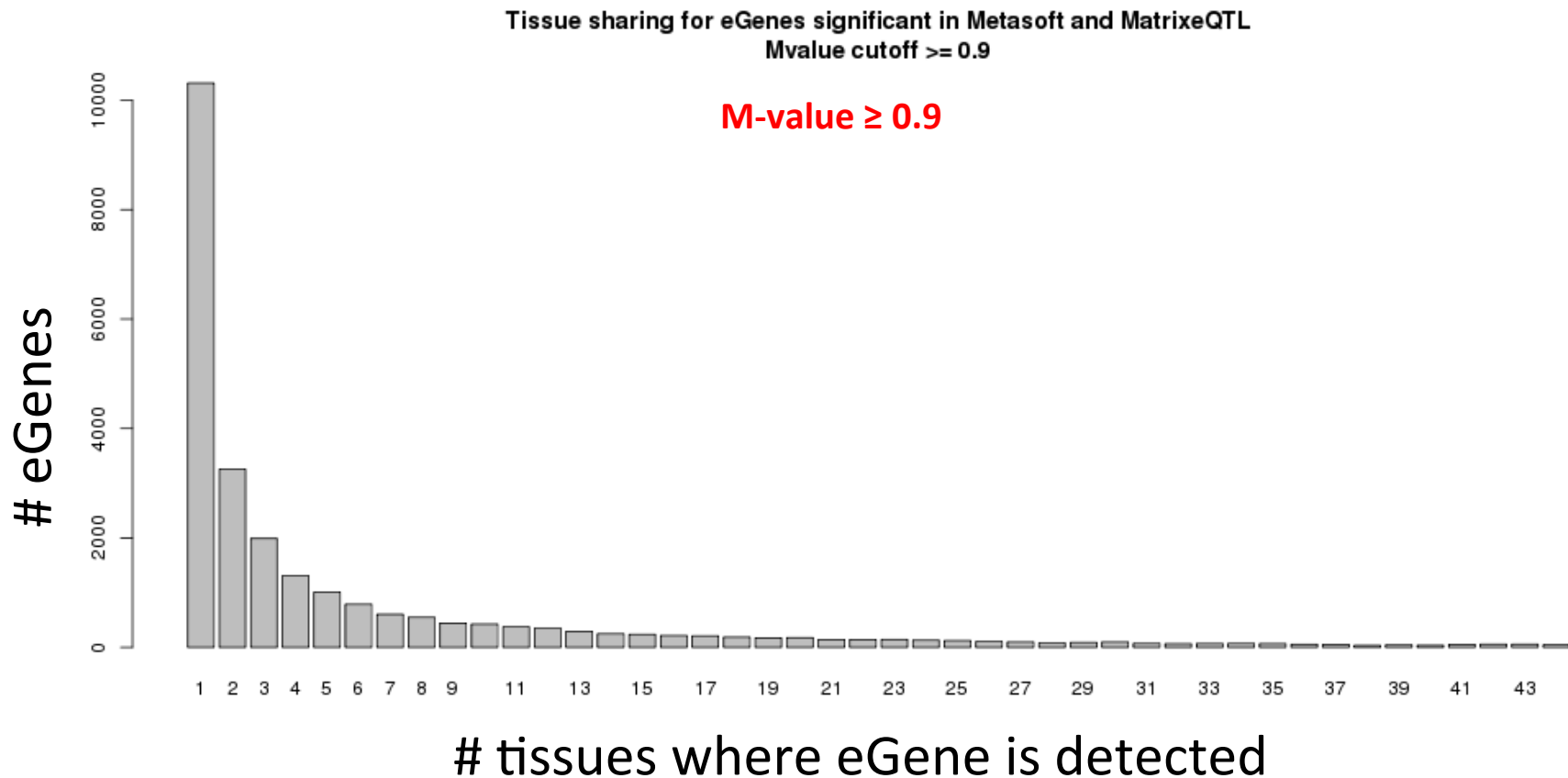| Across 44 tissues	Pi1	Spearman's correlation coefficent
Mean	0.965	0.678
Median	0.973	0.689

Pi1 = estimated true positive rate of significant best eQTLs per eGene from Metasoft (M-value>=0.9) using nominal eQTL SNP-gene pair p-values from Matrix eQTL

Hailei Zhang, Broad Institute

Mean r2 between best eQTL per eGene from Metasoft and Matrix eQTL for overlapping eGenes

tissue.type	mean	median
Skin_Sun_Exposed_Lower_leg	0.54306544	0.528411
Thyroid	0.5419832	0.522761
Testis	0.53882191	0.522522
Esophagus_Muscularis	0.53122901	0.513193
Nerve_Tibial	0.53544066	0.512112
Lung	0.52767276	0.498639
Skin_Not_Sun_Exposed_Suprapubic	0.51944261	0.498
Esophagus_Mucosa	0.51801009	0.494828
Colon_Transverse	0.52473138	0.494295
Cells_Transformed_fibroblasts	0.52220507	0.49386
Cells_EBV-transformed_lymphocytes	0.51414547	0.491133
Adipose_Subcutaneous	0.52608813	0.489324
Artery_Tibial	0.51556262	0.4856575
Breast_Mammary_Tissue	0.51658916	0.4848525
Brain_Cerebellum	0.51338563	0.48121
Heart_Left_Ventricle	0.51716558	0.480633
Artery_Aorta	0.50813268	0.479587
Adipose_Visceral_Omentum	0.50488581	0.47824
Brain_Cerebellar_Hemisphere	0.50399823	0.477549
Heart_Atrial_Appendage	0.51288569	0.477249
Whole_Blood	0.51503176	0.4766975
Muscle_Skeletal	0.51833658	0.47601
Stomach	0.5047372	0.468371
Colon_Sigmoid	0.4974992	0.4677605
Brain_Frontal_Cortex_BA9	0.50690969	0.465665
Brain_Caudate_basal_ganglia	0.49719254	0.462684
Pancreas	0.50608816	0.4621925
Brain_Nucleus_accumbens_basal_ganglia	0.49786436	0.460368
Artery_Coronary	0.4944016	0.459437
Brain_Cortex	0.49466335	0.4484205
Adrenal_Gland	0.49434613	0.446793
Ovary	0.49290756	0.443733
Esophagus_Gastroesophageal_Junction	0.48039371	0.431709
Brain_Hypothalamus	0.46997375	0.426383
Prostate	0.47522125	0.425377
Brain_Putamen_basal_ganglia	0.46119759	0.421678
Pituitary	0.47738741	0.420397
Brain_Anterior_cingulate_cortex_BA24	0.47277781	0.419519
Spleen	0.47133224	0.4186315
Small_Intestine_Terminal_Ileum	0.46370793	0.41397
Brain_Hippocampus	0.44784208	0.4138075
Liver	0.46052913	0.406497
Uterus	0.46417347	0.396866
Vagina	0.43693385	0.385981

Hailei Zhang, Broad Institute

eGene sharing across 44 GTEx tissues for eGenes significant in both Matrix eQTL and Metasoft

Tissue sharing for eGenes significant in Metasoft and MatrixeQTL
Mvalue cutoff >= 0.9

M-value ≥ 0.9

eGenes

tissues where eGene is detected

Hailei Zhang, Broad Institute

Best eQTL per eGene consistency between Matrix eQTL and Metasoft (M-value ≥ 0.9) for tissue-specific and ubiquitous eGenes
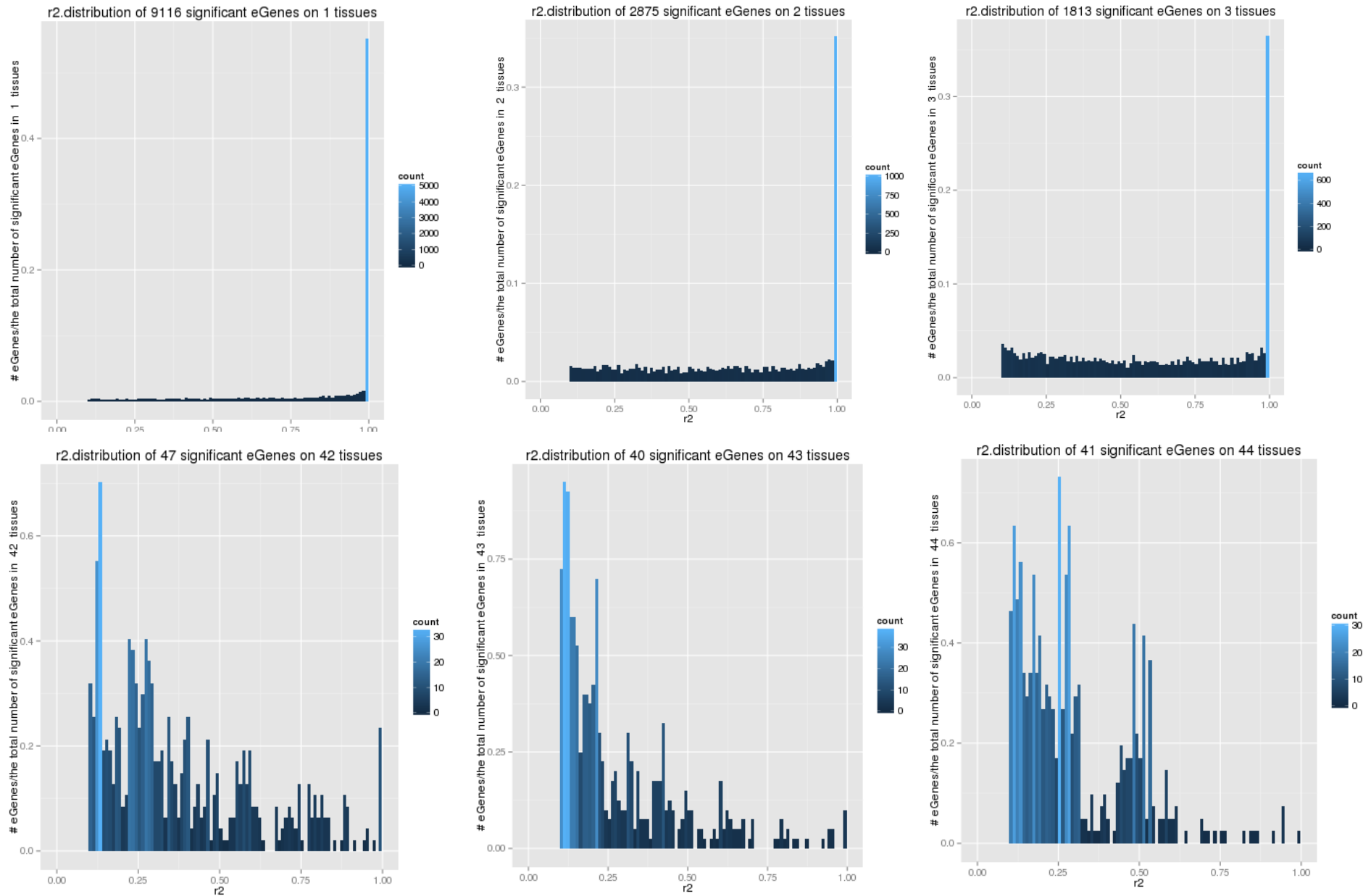
Summary and Next Steps

- The multi-tissue eQTL method Metasoft significantly increases power of eGene detection compared to single tissue analysis (Matrix eQTL) (on average 55-65% more eGenes).
- The overlap between best eQTL per eGene between methods is low. Did we see this also with the UC and UNC multi-tissue methods in the Pilot phase?
- The best eQTL per eGene is more likely to be pointing to the same causal haplotype in more tissue-specific eQTLs vs. ubiquitous eQTLs.
- We need to inspect the consistency of direction of effect of common eQTLs across tissues