Genetic Discovery Summary: 151 Genome-Wide Significant Loci

Table 1: Loci Summary by Signature

Table 1: Genome-wide significant loci identified across 15 disease signatures with genetic signal. For Signature 5, novelty analysis was performed on all 127 genome-wide significant SNPs in the regions; other signatures show independent lead loci.

Signature	Clinical Domain	Total	Unique*	Example Genes
SIG5	Cardiovascular/Lipid	127	10	PCSK9, LPA, LDLR, APOB, APOE, PHACTR1, TRIB1, CELSR2, CDKN2B-AS1
SIG17	GI/Colorectal	19	—	GATA3, BDNF-AS, COLCA1, SMAD3, SMAD7, GREM1, BMP5
SIG7	Hypertension/Vascular	14	—	SH2B3, FGF5, IRF1, CNNM2, HLA-DQB1, KCNK3
SIG19	Skin/Pigmentation	11	—	MC1R, TYR, SLC45A2, CASP8, TOX3, KRT5
SIG1	Musculoskeletal	8	—	GDF5, PRKG1, LTBP1, TGFA, COL27A1
SIG0	Heart Failure/Arrhythmia	6	—	PITX2, LPA, R3HDML, MAP3K7CL
SIG2	Upper GI/GERD	6	—	CCKBR, NCAM1, FOXP1, CRLF1
SIG3	Vascular/Coagulation	6	—	F5, MSRB3, WNT7B, PLCD1
SIG10	Ophthalmologic	6	—	CFH, HTRA1, CDKN2B-AS1, ME3
SIG13	Male Urogenital	5	—	NOTCH4, CLPTM1L, SLC24A4
SIG14	Pulmonary/COPD	4	—	MUC5B, HYKK, CDH23, HHIP-AS1
SIG15	Diabetes/Metabolic	4	—	TCF7L2, HNF1B, UBE2E2, BTNL2
SIG8	Female Reproductive	3	—	TERT, GREB1, EEFSEC
SIG16	Neurodegeneration/Sepsis	1	—	APOE
SIG18	Gallbladder/Biliary	1	—	ABCG8
SIG9	Back Pain/Spine	1	—	NFU1
TOTAL	**Across 15 Signatures**	**151**	**10†**	

*Unique = Not present in any constituent trait GWAS within 1MB windows. Only systematically assessed for Signature 5 (cardiovascular).

†Empirically determined for Signature 5 only. Other signatures lack constituent trait overlap data for novelty assessment.

Signature 5 Novelty Analysis

Key Finding: When comparing Signature 5 (cardiovascular) SNPs to constituent trait GWAS for Angina, Coronary Atherosclerosis, Hypercholesterolemia, Myocardial Infarction, Acute IHD, and Chronic IHD (using 1MB windows):

- **10 of 127 SNPs (7.9%) are unique to Signature 5** — not found in any constituent trait GWAS

- **117 SNPs (92.1%) overlap** with constituent traits, validating biological coherence
- **35 SNPs are highly pleiotropic** — associated with 4+ constituent cardiovascular traits

The 10 Unique Signature 5 Loci (highlighted in green in Table 2):

1. rs6687726 (**IL6R**) — Interleukin-6 receptor, inflammatory pathway
2. rs2509121 (**HYOU1**) — Hypoxia up-regulated protein, ER stress response
3. rs4760278 (**R3HDM2**) — RNA binding protein
4. rs1532085 (**LIPC**) — Hepatic lipase, HDL metabolism
5. rs7168222 (**NR2F2-AS1**) — Nuclear receptor antisense RNA
6. rs35039495 (**PLCG2**) — Phospholipase C gamma 2, platelet activation
7. rs8121509 (**OPRL1**) — Opioid receptor, pain/stress signaling
8. rs1499813 (**FNDC3B**) — Fibronectin domain, insulin signaling
9. 4:96088139 (**UNC5C**) — Netrin receptor, axon guidance
10. rs4732365 (**C7orf55**) — Unknown function

Table 2: Representative Genetic Associations with Biological Validation

Table 2: Selected loci demonstrating biological coherence of signature-based GWAS. Green highlighting indicates loci unique to Signature 5 (not in any constituent trait GWAS).

Signature	SNP	Gene	P-value	Biological Context
SIGNATURE 5: Cardiovascular/Lipid-Mediated Disease				
56 total loci; 10 unique to Sig5 (highlighted in green), 46 overlap with constituent traits				
SIG5	rs10455872	LPA	2.8×10^{-130}	Lipoprotein(a), strongest CAD association
SIG5	rs1333042	CDKN2B-AS1	9.0×10^{-100}	9p21 locus, CAD, MI, glaucoma (pleiotropy)
SIG5	rs7412	APOE	1.8×10^{-59}	APOE-ε2 allele, lipids, Alzheimer's
SIG5	rs12740374	CELSR2	1.7×10^{-42}	Cadherin EGF LAG seven-pass, LDL-C
SIG5	rs138294113	LDLR	5.9×10^{-39}	LDL receptor, familial hypercholesterolemia
SIG5	rs9349379	PHACTR1	6.5×10^{-30}	Phosphatase and actin regulator, CAD
SIG5	rs11591147	PCSK9	7.5×10^{-25}	Proprotein convertase, LDL-C, CAD therapeutic target
SIG5	rs2954031	TRIB1	2.5×10^{-15}	Tribbles pseudokinase 1, lipid metabolism
SIG5	rs2119690	LPL	1.8×10^{-10}	Lipoprotein lipase, triglyceride hydrolysis
SIG5	**rs1499813**	**FNDC3B**	$\mathbf{1.1 \times 10^{-10}}$	**UNIQUE: Fibronectin domain, adipogenesis, insulin signaling**
SIG5	**rs2509121**	**HYOU1**	$\mathbf{1.0 \times 10^{-9}}$	**UNIQUE: Hypoxia up-regulated protein, ER stress**
SIG5	**rs4732365**	**C7orf55**	$\mathbf{1.1 \times 10^{-9}}$	**UNIQUE: Chromosome 7 ORF, unknown function**
SIG5	**rs8121509**	**OPRL1**	$\mathbf{3.7 \times 10^{-9}}$	**UNIQUE: Opioid receptor, stress/pain signaling**
SIG5	**rs6687726**	**IL6R**	$\mathbf{1.3 \times 10^{-8}}$	**UNIQUE: Interleukin-6 receptor, systemic inflammation**
SIG5	**rs35039495**	**PLCG2**	$\mathbf{2.6 \times 10^{-8}}$	**UNIQUE: Phospholipase C, platelet activation, immunity**

Continued on next page

Signature	SNP	Gene	P-value	Biological Context
SIG5	**rs7168222**	**NR2F2-AS1**	$\mathbf{2.7 \times 10^{-8}}$	**UNIQUE: Nuclear receptor antisense RNA**
SIG5	**rs4760278**	**R3HDM2**	$\mathbf{3.2 \times 10^{-8}}$	**UNIQUE: R3H domain containing 2, RNA binding**
SIG5	**4:96088139**	**UNC5C**	$\mathbf{3.7 \times 10^{-8}}$	**UNIQUE: Unc-5 netrin receptor, axon guidance**
SIG5	**rs1532085**	**LIPC**	$\mathbf{3.8 \times 10^{-8}}$	**UNIQUE: Hepatic lipase, HDL metabolism**
SIGNATURE 15: Diabetes and Metabolic Disease				
SIG15	rs7903146	TCF7L2	5.1×10^{-28}	Transcription factor 7-like 2, strongest T2D association
SIG15	rs11657964	HNF1B	2.8×10^{-8}	Hepatocyte nuclear factor, MODY5, renal cysts
SIG15	rs1496653	UBE2E2	1.4×10^{-8}	Near diabetes-associated locus
SIG15	rs3806155	BTNL2	3.3×10^{-10}	MHC region, immune-metabolic interface
SIGNATURE 10: Age-Related Ophthalmologic Disease				
SIG10	rs33944729	CFH	1.1×10^{-11}	Complement factor H, macular degeneration
SIG10	rs58077526	HTRA1	6.1×10^{-17}	Serine peptidase, age-related macular degeneration
SIG10	rs7866783	CDKN2B-AS1	1.9×10^{-8}	9p21, glaucoma, retinal disease (also in SIG5, SIG17, SIG19)
SIGNATURE 0: Heart Failure and Arrhythmia				
SIG0	rs6843082	PITX2	1.6×10^{-17}	Paired-like homeodomain TF, atrial fibrillation
SIG0	rs74617384	LPA	5.0×10^{-10}	Lipoprotein(a), shared with SIG5 (pleiotropy)
SIGNATURE 17: Gastrointestinal/Colorectal				
SIG17	rs12777423	GATA3	3.9×10^{-8}	Transcription factor, GI development
SIG17	rs17309874	BDNF-AS	3.3×10^{-11}	Brain-derived neurotrophic factor, gut-brain axis
SIG17	rs7103178	COLCA1	8.5×10^{-9}	Colorectal cancer associated 1/2
SIG17	rs4939567	SMAD7	5.4×10^{-16}	TGF-β signaling, colorectal cancer
SIG17	rs58658771	GREM1	1.3×10^{-18}	BMP antagonist, colorectal cancer
SIGNATURE 14: Pulmonary/COPD				
SIG14	rs35705950	MUC5B	1.3×10^{-8}	Mucin 5B, idiopathic pulmonary fibrosis
SIG14	rs72738786	HYKK	4.3×10^{-19}	Near CHRNA5-CHRNA3-CHRNB4, lung function
SIG14	rs6537293	HHIP-AS1	4.3×10^{-8}	Hedgehog interacting protein, COPD, lung function
SIGNATURE 13: Male Urogenital Disease				
SIG13	rs460073	CLPTM1L	1.7×10^{-14}	TERT-CLPTM1L region, prostate cancer, bladder cancer
SIG13	rs2071278	NOTCH4	3.3×10^{-10}	MHC region, prostate associations (also in SIG3, SIG5, SIG7, SIG15, SIG17)
SIGNATURE 7: Hypertension and Vascular Regulation				
SIG7	rs3184504	SH2B3	1.1×10^{-13}	SH2B adaptor protein 3, blood pressure, inflammation

Continued on next page

Signature	SNP	Gene	P-value	Biological Context
SIG7	rs12509595	FGF5	1.6×10^{-10}	Fibroblast growth factor 5, blood pressure
SIG7	rs11242115	IRF1	1.3×10^{-8}	Interferon regulatory factor, immune-vascular
SIG7	rs9275218	HLA-DQB1	5.5×10^{-17}	MHC class II (7 signatures total—maximum pleiotropy)
SIGNATURE 16: Neurodegeneration/Acute Illness				
SIG16	rs429358	APOE	9.9×10^{-14}	APOE-ε4, Alzheimer's, cognitive decline, sepsis
SIGNATURE 18: Gallbladder and Biliary Disease				
SIG18	rs11887534	ABCG8	1.3×10^{-55}	ATP-binding cassette, cholesterol transport, gallstones
SIGNATURE 19: Skin and Pigmentation				
SIG19	rs1126809	TYR	4.7×10^{-14}	Tyrosinase, melanin synthesis
SIG19	rs16891982	SLC45A2	3.5×10^{-13}	Solute carrier, pigmentation
SIG19	rs1805007	MC1R	1.5×10^{-12}	Melanocortin 1 receptor, red hair, skin cancer

Key Findings

1. **151 genome-wide significant loci** identified across 15 disease signatures
2. **Biological coherence:** Nearly all loci map to genes with established roles in signature-relevant pathways:
   - SIG5 (cardiovascular): Lipid genes (PCSK9, LPA, LDLR, APOB, APOE, CELSR2, LPL, TRIB1)
   - SIG15 (diabetes): Glucose regulation (TCF7L2, HNF1B)
   - SIG10 (ophthalmologic): Macular degeneration genes (CFH, HTRA1)
   - SIG14 (pulmonary): Lung function and fibrosis (MUC5B, HHIP)
   - SIG19 (skin): Pigmentation genes (TYR, MC1R, SLC45A2)
3. **Novel discoveries:** 10 loci unique to Signature 5 not found in any constituent cardiovascular trait GWAS, including biologically relevant genes:
   - **IL6R** (inflammation), **PLCG2** (platelet activation)
   - **FNDC3B** (insulin signaling), **LIPC** (hepatic lipase, HDL metabolism)
   - **HYOU1** (hypoxia/ER stress), **OPRL1** (stress response)
4. **Pleiotropy revealed:** 49 loci (32.5%) affect multiple signatures:
   - **HLA-DQB1**: 7 signatures (maximum pleiotropy)
   - **CDKN2B-AS1**: SIG5, SIG10, SIG17, SIG19 (cardiovascular, ophthalmologic, GI, skin)
   - **LPA**: SIG0 and SIG5 (heart failure and cardiovascular)
   - **NOTCH4**: 6 signatures (MHC region, autoimmune-metabolic interface)
5. **Joint modeling power:** For Signature 5, 92.1% of loci overlap with constituent traits (validating biology), while 7.9% are unique discoveries—demonstrating both validation and enhanced discovery
6. **Highly pleiotropic core:** 35 Signature 5 loci associate with 4+ cardiovascular conditions, revealing genetic architecture of disease COMBINATIONS not visible in single-trait GWAS
7. **Phenome-wide coverage:** Genetic discoveries span cardiovascular, metabolic, pulmonary, GI, ophthalmologic, neurologic, reproductive, and dermatologic domains

Table 3: rs1532085 near LIPC demonstrates signature-based discovery of biologically meaningful loci

Analysis	**P-value**	**Result**
Angina (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
MI (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
Hypercholesterolemia (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
Coronary atherosclerosis (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
Acute IHD (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
Chronic IHD (individual GWAS)	$p > 5 \times 10^{-8}$	Not significant
Signature 5 (joint GWAS)	$\mathbf{3.8 \times 10^{-8}}$	**Genome-wide significant**

Detailed Example: rs1532085 (LIPC) — A Unique Signature 5 Discovery

Biological interpretation: LIPC encodes hepatic lipase, a critical enzyme in HDL (high-density lipoprotein) metabolism and reverse cholesterol transport. Hepatic lipase plays multiple roles in cardiovascular disease:

- Hydrolysis of triglycerides and phospholipids in HDL and LDL particles
- Regulation of HDL particle size and composition
- Modulation of hepatic lipoprotein uptake
- Effects on remnant lipoprotein clearance and atherosclerosis

This variant affects cardiovascular disease through *distributed pleiotropic effects* across multiple manifestations (angina, MI, atherosclerosis, hypercholesterolemia)—each individually too weak to detect in single-trait GWAS, but collectively significant when modeled jointly through the cardiovascular signature.

This exemplifies signature-based discovery: The variant is at a well-established lipid metabolism gene (LIPC) but only detectable through joint modeling of related cardiovascular disease phenotypes. Individual cardiovascular endpoints dilute the signal, while the signature captures the aggregate lipid-mediated cardiovascular effect.

Interpretation for Reviewers

Novelty Framework:

- **7.9% completely novel** (10/127 Sig5 SNPs) — not in constituent trait GWAS
- **92.1% validate known biology** — overlap confirms biological coherence
- **27.6% highly pleiotropic** (35/127 Sig5 SNPs) — affect 4+ cardiovascular traits

The 10 unique loci include genes with clear biological relevance (IL6R, PLCG2, FNDC3B, LIPC) that connect cardiovascular disease to inflammation, platelet function, insulin signaling, and HDL metabolism—pathways that are biologically well-established but missed when analyzing traditional cardiovascular endpoints in isolation.

This is not statistical artifact. This is biological signal revealing disease processes and comorbidity mechanisms beyond single-disease risk factors.