❖ *Author's Choice*

Systems-based approaches for investigation of inter-tissue communication🅢

Marcus M. Seldin* and Aldons J. Lusis[1],*,†,§

Departments of Medicine,* Human Genetics,† and Microbiology, Immunology, and Molecular Genetics,§ University of California, Los Angeles, Los Angeles, CA 90095

Abstract Secreted proteins serve as crucial mediators of many physiology processes, and beginning with the discovery of insulin, studies have revealed numerous context-specific regulatory networks across various cell types. Here, we review "omics" approaches to deconvolute the complex milieu of proteins that are released from the cell. We emphasize a novel "systems genetics" approach our laboratory has developed to investigate mechanisms of tissue-tissue communication using population-based datasets.🅙 Finally, we highlight potential future directions for these studies, discuss several caveats, and propose new ways to investigate modes of endocrine communication.—Seldin, M. M., and A. J. Lusis. Systems-based approaches for investigation of inter-tissue communication. *J. Lipid Res.* 2019. **60: 450–455.**

Supplementary key words cytokines • diabetes • genetics • hormones • receptors/hormone
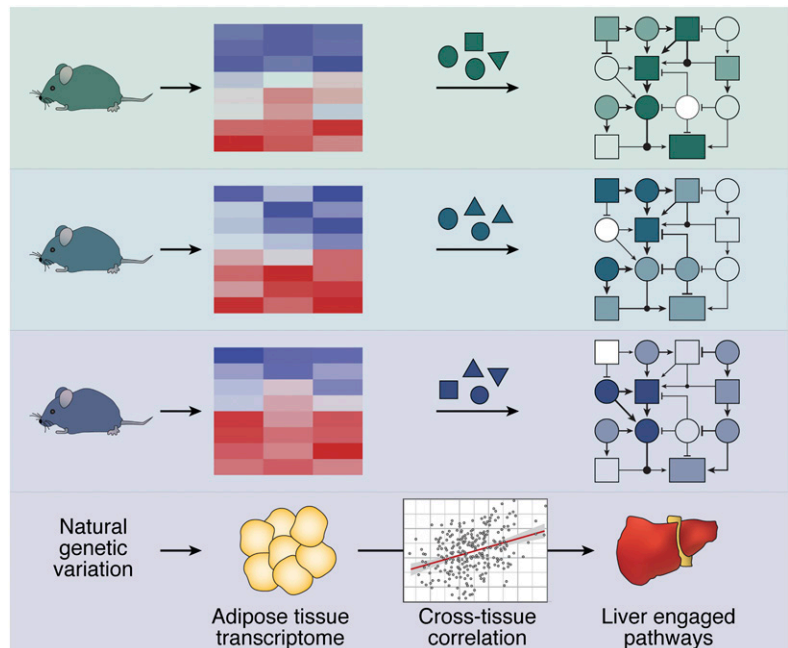
SYSTEMS GENETICS

Systems genetics, an approach that involves the integration of clinical/physiological and omics level data in natural populations, is proving particularly useful for the analysis of complex traits (1). For example, quantification of global mRNA levels for a particular tissue in a population can be used to identify common genetic variants that control transcription, termed expression quantitative trait loci (eQTLs). These can be related to clinical traits by co-mapping, correlation or statistical modeling to help identify candidate disease mechanisms. In this review, we discuss the application of systems genetics approaches to identify novel endocrine factors mediating tissue-tissue cross-talk.

ABUNDANCE OF SECRETED PROTEINS IN MAMMALS

Secreted proteins constitute a significant fraction (roughly 15–20%) of the coding genome (2). These serve a variety of functions, including the relay of signals to neighboring cells, a process conserved from bacteria to humans. Since the discovery of insulin over a century ago, many elegant studies have related circulating proteins to regulatory and homeostatic functions, and these proteins have been found to be involved in many common human diseases. Research

---

on molecules, such as leptin, adiponectin, ghrelin, and fibroblast growth factor family members, has driven our understanding of endocrine metabolism and guided therapeutic development for diseases such as obesity and type 2 diabetes. At present, hundreds of additional secreted proteins have been described to regulate diverse cellular processes, elucidated largely through gain- or loss-of-function studies. As a result, various clinical trials have shown promise in modulating endocrine factors' levels in circulation through antibody-based inhibition or direct administration of recombinant proteins. Understanding of the functions of the vast majority of secreted proteins, however, is still lacking, presenting an exciting avenue for scientific discovery and therapeutic development.

ASSAYS AND METHODS TO INTERROGATE CELLULAR SECRETOME

The past decade has seen an immense increase in the generation of large omics datasets aimed at generating a more global picture of cellular states. Specifically, advances in efficiency, cost reduction, and analytical approaches to interpret transcriptomic (expression arrays or RNA-sequencing) and/or proteomic (multiplexed antibody/ aptamer or mass spectrometry-based quantification) have further helped to guide our understanding of the cellular "secretome". Gene expression or quantitative protein datasets have been filtered for either the predicted or annotated presence of secretion signals to gain insight into context-specific regulation of the secretome at the transcript level (3, 4). Likewise, proteomics can be performed directly on culture media (5–7) or plasma (8–10). Several recent studies have identified novel genetic mechanisms of the circulating proteome that show enrichment in disease (11, 12). Some relative advantages and disadvantages of each strategy have been described in greater detail (13–15). A recent study performed cytokine profiling in blood cells under a variety of pathologic stimuli from 500 healthy individuals. By pairing these observations with genetic data, the authors provided a snapshot of genetic regulation of the secretome under normal or perturbed conditions (16). These generalized approaches have substantially contributed to our understanding of cellular regulation of secretory pathways and revolutionized biomarker discovery. Given that these observations are mostly correlative in nature, it can be quite difficult to infer the physiologic functions of the secreted molecules identified and determine whether they are causal or reactive with respect to a physiological or disease process.

DECONVOLUTION OF ENDOCRINE INTERACTIONS USING NATURAL VARIATION

Combined analyses of omics data across populations have been instrumental for inference of pathways regulating diverse cellular and tissue functions. One such example includes the STARNET (Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task) population, where genotyping was combined with RNA-sequencing from several

tissues to guide disease inference (17). This initial study also noted strong enrichment of human genome-wide association study disease loci in networks that persisted across tissues, suggesting an important role for signaling between multiple tissues in disease progression. Further, the notion that a gene network model across tissues could explain genetic differences between susceptibility to pancreatic islet dysfunction was put forth by comparing two strains in mice (18). In 2016, Long et al. (19) constructed a coexpression network of heart and blood transcripts from the GTeX (Genotype-Tissue Expression Project) dataset, noting that a secreted protein, dipeptidyl peptidase 4, could act as a link mediating signaling between the tissues (19). These studies and several others highlight the potential of comparing gene expression across tissues from genetically diverse individuals.

We recently developed a framework that utilizes natural variation and multi-tissue datasets to identify modes of endocrine communication, termed Quantitative Endocrine Network Interaction Estimation (QENIE) (20). The basic concept is simply to search for genes that strongly correlate with transcriptome measures across tissues within a population. For our study, we used the hybrid mouse diversity panel (HMDP), a collection of about 100 "classical" as well as recombinant inbred strains of mice. Using global transcriptomic data from several tissues, we found that correlation structure between tissues predicted many known interactions, such as adiponectin acting from adipose tissue to liver. The intuition underling this prioritization is that the hepatic pathways and functions engaged by adiponectin would be apparent through correlations with the gene in adipose tissue (**Fig. 1**). For genes exhibiting such correlation structure, we hypothesized that we could use pathway enrichment analysis and integration with available clinical trait data to generate testable hypotheses as to function. Using this approach, we identified and characterized several novel genes that mediate signaling and affect functions across tissues. For example, we showed that adipose-derived lipocalin-5 is an enhancer of skeletal muscle respiration, and that liver-derived Notum is a promoter of adipose tissue thermogenesis. We also provided evidence that inter-alpha-trypsin inhibitor heavy chain family member 5 suppresses heart starvation response, that SPARC-related modular calcium binding 1 regulates cardiac cell proliferation, and that pro-platelet basic protein in enhances vascular inflammation. This approach is generalizable in that it can be easily applied to mouse or human or other populations that have been subjected to appropriate transcriptomic or proteomic analyses.

FUTURE APPROACHES TO INVESTIGATE TISSUE-TISSUE COMMUNICATION

Our initial framework begins by ranking interactions by global patterns of correlation, which biases predictions toward larger pathways of target tissue genes. An alternative approach would be to focus on specific pathways or correlated modules of genes in a given target tissue, then rank origin tissue secreted factors based exclusively on these genes. This is illustrated in **Fig. 2A** between liver and adipose
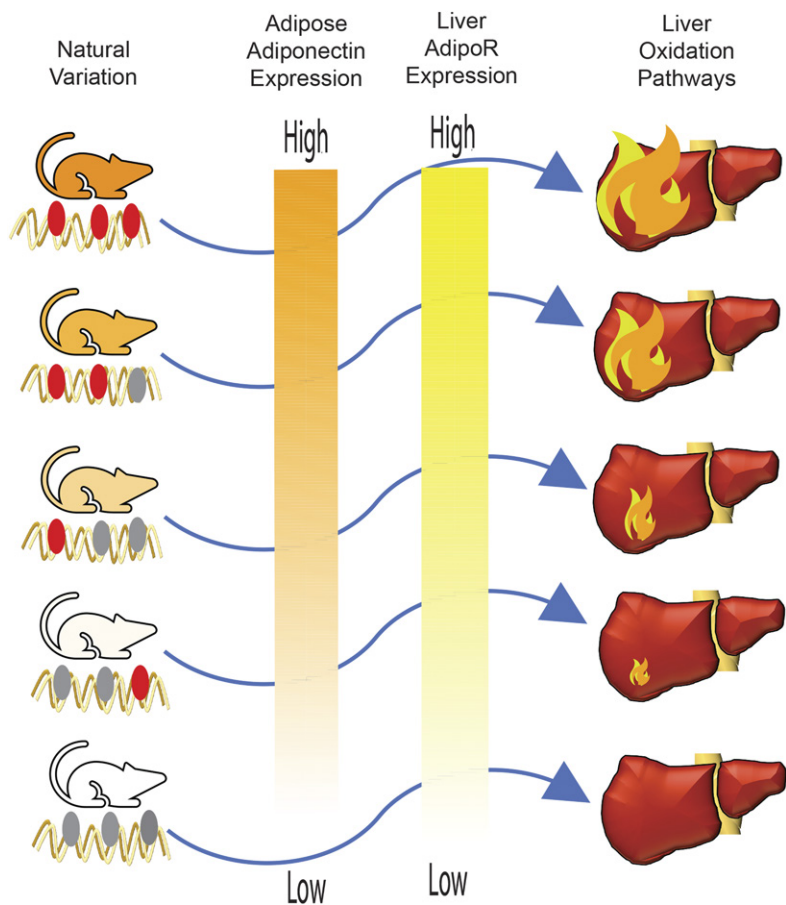
Fig. 1. Intuition for QENIE. Illustration of the intuition for QENIE approach from adipose tissue to liver. Natural variation drives different levels of adipose tissue expression of an adiponectin (left). Given that adiponectin engages hepatic oxidation pathways through its receptor, the abundance of AdipoRs and consequent oxidative pathways (right) will show similar quantitative levels to that of adiponectin in fat.

tissue, where genes predicted to target individual modules are indicated. Here, an unbiased approach to network construction, weighted gene correlated network analysis (WGCNA), elucidated several different modules of correlated genes. These individual modules were enriched for very different pathways, which in turn were correlated with different secreted proteins in the reciprocal tissue. It is important to note that regardless of how the interactions are constructed, experimental validation is required, as it is impossible to eliminate all confounding factors. Along these lines, a different approach could utilize regulatory genetic variation to drive correlations. For example, focusing on a *cis*-eQTL for a given secreted protein, which also maps in *trans* to a target tissue pathway, could provide evidence of causal inference. However, one caveat with this approach would be the exclusion of *trans*-regulated genes coding for secreted proteins. Given that, generally, *cis*-regulation tends to explain a small fraction of variation in gene expression as well as trait association (21–23) and that several well-described endocrine factors, such as adiponectin, leptin, resistin, and angiopoietin-like 4, are not regulated in *cis* in adipose tissue within the HMDP, we chose to focus on a correlation-based approach.

Perhaps the most obvious potential future application of QENIE or similar approaches is to integrate quantitative proteomics data. This can be accomplished by either *1)* using tissue-specific proteomics data to make similar cross-tissue predictions or *2)* pairing either tissue gene expression and/or proteomics correlations with quantitative plasma protein data to refine predictions. Here, one could simply correlate pathways in tissues with adipose and/or plasma levels of a secreted factor. For example, the heterotrimeric structure of secreted adiponectin has been shown to play an integral role for its function, which could be measured in the HMDP and presumably correlate even more strongly with liver oxidative gene expression (24–26). Given that levels of protein in a specific tissue or plasma are much closer to the mode of action, these analyses are potentially more powerful than gene expression. Similar to reliance on correlation with transcript levels, integration of quantitative trait loci regulating proteins would further aid causal inference. This would be particularly valuable in instances where endocrine regulation of a specific pathway is being driven by posttranscriptional control of a circulating factor. Several notable studies have described distinct patterns of *cis*-regulated proteins, which cannot be detected at the transcript level (27). One constraint of using proteomics is the limitation of depth, as RNA-sequencing tends to reliably detect around 4- to 6-fold more coding genes when compared with mass-spectrometry proteomics.

As mentioned above, one important constraint in the current QENIE pipeline is the exclusive reliance on correlation to predict interactions. While we feel that this enables the most potential for discovery of endocrine circuits, this approach is subject to spurious correlations and therefore, requires substantially more experimental validation than other predicted (e.g., overlapping *cis*-eQTLs with modules in other tissues). Given that many correlations are
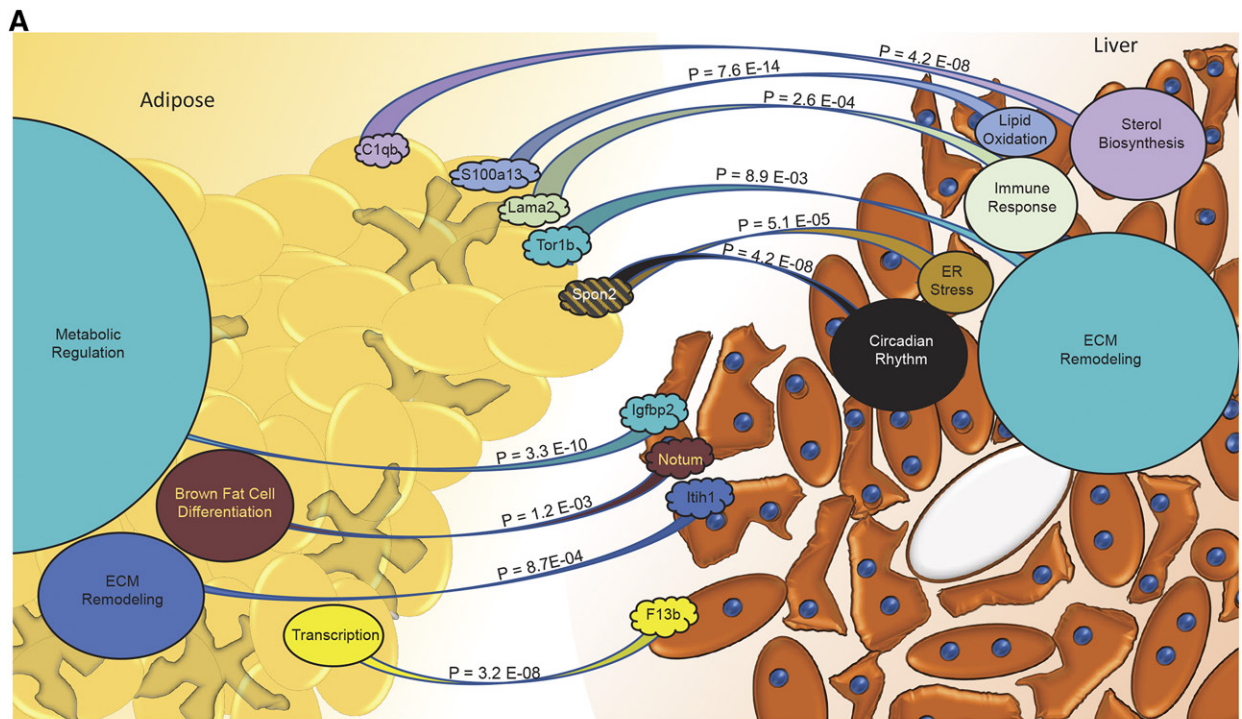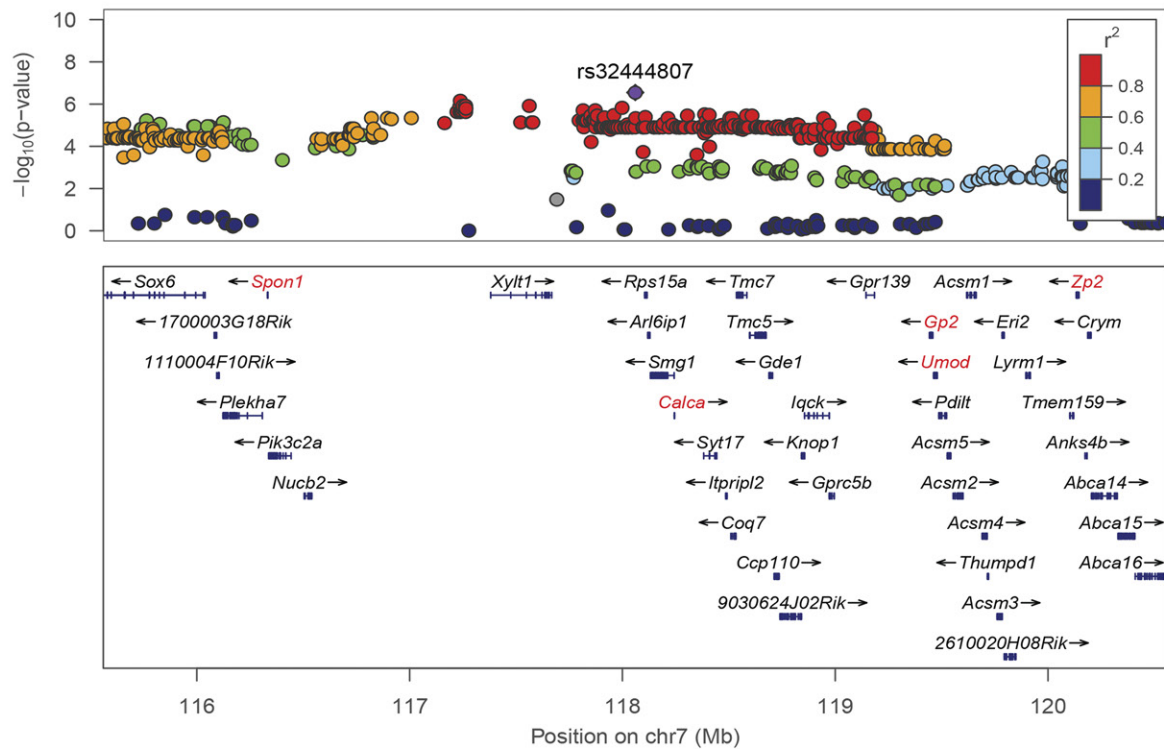
Fig. 2. Pathway-based interactions between liver and adipose tissue. A: Each pathway was constructed from weighted gene correlated network analysis (WGCNA), where the size of the circle indicates the number of genes in each respective module. The strongest predicted genes encoding secreted proteins in the neighboring tissue for each module is shown, with the *P* value based on permutation testing. B: Locuszoom plot showing a genome-wide significant region on chromosome 7 for the adipose tissue metabolic regulation module eigengene. Factors that are secreted are highlighted in red.

reactive in nature, many cross-tissue predictions follow the same pattern. For example, if both the expression of a secreted factor in one tissue and the global pathway in another are regulated by insulin, correlation would show

a direct relationship, which may be spurious. One could regress these correlations on plasma insulin to tease out these effects; but given the complexity of these processes, it is quite difficult to pinpoint potential sources of spurious

relationships. Consistent with these considerations, we have observed an approximate 15–25% "validation rate" of such interactions, depending on dataset used. For this reason and others, it is crucial to follow predictions with experimental validation.

An alternative to identification of endocrine regulation through co-mapping of *cis*-eQTLs or *cis*-pQTLs with pathways in neighboring tissues, which would likely improve causal inference from QENIE, would be to simply evaluate genes near a locus for the pathway of interest. Here, the approach would be to map modules of gene expression and look for potential secreted factors under the locus. To demonstrate the utility of such approaches, genome-wide association was performed on the eigengene for the adipose tissue metabolic regulation module, where several endocrine factors surrounding a significant quantitative trait locus are highlighted in red (**Fig. 2B**). The appeal here would be that there is little reliance on whether a *cis* variant is acting at the level of gene, protein, or secretion; however, without integration of additional data uncovering which protein is driving the module and what tissue is the primary contributor to the observed effect would be difficult.

Additionally, one could integrate this approach with metabolomics datasets. In our initial study, we compared the average level of correlation across tissues for genes coding for secreted proteins versus all other genes, where most tissue-tissue pairs showed stronger correlation enrichment for secreted proteins. We noted several exceptions to this observation. For example, nonsecreted genes from skeletal muscle correlated more strongly than those of secreted proteins across the adipose tissue transcriptome. One possible explanation for this observation is that these genes are involved in the production of a metabolite that relays a signal between the two tissues. Previous studies have described a strong genetic influence of the plasma metabolome (28–31) and, given that many of these molecules are known to relay signals between tissues, one could expect to uncover some new functions by integrating the data in a similar manner.

An area of research that has gained considerable recent attention is the investigation of circulating microRNAs as signaling mechanisms between tissues. These microRNAs can be found either packaged in lipoprotein particles (32) or exosomes (33), or bound directly by protein complexes (34), and can impact a variety of tissues. These microRNAs could also serve as the basis for analyses of communication across tissues, as opposed to proteins, which are secreted.

Beyond modification of the pipeline or application to other datasets, there are some intriguing future directions for some of the endocrine factors we screened and identified. For example, identification of the upstream physiologic parameters that regulate expression and secretion of the proteins will be crucial to uncovering their conserved roles. Whether genes such as lipocalin-5 and Notum are regulated by fasting/feeding, exercise, or cold exposure could guide future studies. Further, knowing the receptors by which these proteins act would be highly informative as to their functional and tissue-specific impacts. While the receptor by which lipocalin-5 signals is not known, Notum

has been shown to be a potent inhibitor of Wnt signaling (35). From these known observations, one can begin to ask how Notum links Wnt signaling inhibition to adipose tissue thermogenesis.

In conclusion, systems genetics approaches provide a powerful complement to traditional methods for endocrine discovery. Much remains to be learned because the functional outcomes of hundreds of secreted proteins remain unknown. A major challenge is the understanding of how these novel factors function in homeostasis and contribute to disease.

REFERENCES

1. Civelek, M., and A. J. Lusis. 2014. Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* **15:** 34–48.
2. Uhlén, M., L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, et al. 2015. Proteomics. Tissue-based map of the human proteome. *Science.* **347:** 1260419.
3. Nielsen, H. 2017. Predicting secretory proteins with SignalP. *Methods Mol. Biol.* **1611:** 59–73.
4. Chen, Y., Y. Zhang, Y. Yin, G. Gao, S. Li, Y. Jiang, X. Gu, and J. Luo. 2005. SPD–a web-based secreted protein database. *Nucleic Acids Res.* **33:** D169–D173.
5. Meissner, F., R. A. Scheltema, H. J. Mollenkopf, and M. Mann. 2013. Direct proteomic quantification of the secretome of activated immune cells. *Science.* **340:** 475–478.
6. Kratchmarova, I., D. E. Kalume, B. Blagoev, P. E. Scherer, A. V. Podtelejnikov, H. Molina, P. E. Bickel, J. S. Andersen, M. M. Fernandez, J. Bunkenborg, et al. 2002. A proteomic approach for identification of secreted proteins during the differentiation of 3T3–L1 preadipocytes to adipocytes. *Mol. Cell. Proteomics.* **1:** 213–222.
7. Chan, C. Y., O. Masui, O. Krakovska, V. E. Belozerov, S. Voisin, S. Ghanny, J. Chen, D. Moyez, P. Zhu, K. R. Evans, et al. 2011. Identification of differentially regulated secretome components during skeletal myogenesis. *Mol. Cell. Proteomics.* **10:** M110.004804.
8. Nanjappa, V., J. K. Thomas, A. Marimuthu, B. Muthusamy, A. Radhakrishnan, R. Sharma, A. Ahmad Khan, L. Balakrishnan, N. A. Sahasrabuddhe, S. Kumar, et al. 2014. Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic Acids Res.* **42:** D959–D965.
9. Geyer, P. E., N. A. Kulak, G. Pichler, L. M. Holdt, D. Teupser, and M. Mann. 2016. Plasma proteome profiling to assess human health and disease. *Cell Syst.* **2:** 185–195.
10. Desiere, F., E. W. Deutsch, A. I. Nesvizhskii, P. Mallick, N. L. King, J. K. Eng, A. Aderem, R. Boyle, E. Brunner, S. Donohoe, et al. 2005. Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol.* **6:** R9.
11. Sun, B. B., J. C. Maranville, J. E. Peters, D. Stacey, J. R. Staley, J. Blackshaw, S. Burgess, T. Jiang, E. Paige, P. Surendran, et al. 2018. Genomic atlas of the human plasma proteome. *Nature.* **558:** 73–79.
12. Emilsson, V., M. Ilkov, J. R. Lamb, N. Finkel, E. F. Gudmundsson, R. Pitts, H. Hoover, V. Gudmundsdottir, S. R. Horman, T. Aspelund, et al. 2018. Co-regulatory networks of human serum proteins link genetics to disease. *Science.* **361:** 769–773.
13. Stastna, M., and J. E. Van Eyk. 2012. Investigating the secretome: lessons about the cells that comprise the heart. *Circ Cardiovasc Genet.* **5:** o8–o18.
14. Hathout, Y. 2007. Approaches to the study of the cell secretome. *Expert Rev. Proteomics.* **4:** 239–248.
15. Mukherjee, P., and S. Mani. 2013. Methodologies to decipher the cell secretome. *Biochim. Biophys. Acta.* **1834:** 2226–2232.
16. Li, Y., M. Oosting, S. P. Smeekens, M. Jaeger, R. Aguirre-Gamboa, K. T. T. Le, P. Deelen, I. Ricano-Ponce, T. Schoffelen, A. F. M. Jansen, et ale, L. A. B. Joosten, C. Wijmenga, V. Kumar, and M. G. Netea. 2016. A functional genomics approach to understand variation in cytokine production in humans. *Cell.* **167:** 1099–1110.e14.
17. Franzén, O., R. Ermel, A. Cohain, N. K. Akers, A. Di Narzo, H. A. Talukdar, H. Foroughi-Asl, C. Giambartolomei, J. F. Fullard, K. Sukhavasi, et al. 2016. Cardiometabolic risk loci share downstream

cis- and trans-gene regulation across tissues and diseases. *Science.* **353:** 827–830.

18. Keller, M. P., Y. Choi, P. Wang, D. B. Davis, M. E. Rabaglia, A. T. Oler, D. S. Stapleton, C. Argmann, K. L. Schueler, S. Edwards, et al. 2008. A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. *Genome Res.* **18:** 706–716.

19. Long, Q., C. Argmann, S. M. Houten, T. Huang, S. Peng, Y. Zhao, Z. Tu, G. T. Consortium, and J. Zhu. 2016. Inter-tissue coexpression network analysis reveals DPP4 as an important gene in heart to blood communication. *Genome Med.* **8:** 15.

20. Seldin, M. M., S. Koplev, P. Rajbhandari, L. Vergnes, G. M. Rosenberg, Y. Meng, C. Pan, T. M. N. Phuong, R. Gharakhanian, N. Che, et al. 2018. A strategy for discovery of endocrine interactions with application to whole-body metabolism. *Cell Metab.* **27:** 1138–1155.e6.

21. Grundberg, E., K. S. Small, A. K. Hedman, A. C. Nica, A. Buil, S. Keildson, J. T. Bell, T. P. Yang, E. Meduri, A. Barrett, et al. 2012. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat. Genet.* **44:** 1084–1089.

22. Rockman, M. V., and L. Kruglyak. 2006. Genetics of global gene expression. *Nat. Rev. Genet.* **7:** 862–872.

23. Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, et al. 2009. Finding the missing heritability of complex diseases. *Nature.* **461:** 747–753.

24. Waki, H., T. Yamauchi, J. Kamon, Y. Ito, S. Uchida, S. Kita, K. Hara, Y. Hada, F. Vasseur, P. Froguel, et al. 2003. Impaired multimerization of human adiponectin mutants associated with diabetes. Molecular structure and multimer formation of adiponectin. *J. Biol. Chem.* **278:** 40352–40363.

25. Hug, C., and H. F. Lodish. 2005. The role of the adipocyte hormone adiponectin in cardiovascular disease. *Curr. Opin. Pharmacol.* **5:** 129–134.

26. Liu, M., and F. Liu. 2014. Regulation of adiponectin multimerization, signaling and function. *Best Pract. Res. Clin. Endocrinol. Metab.* **28:** 25–31.

27. Chick, J. M., S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi, D. M. Gatti, N. Raghupathy, K. L. Svenson, G. A. Churchill, and S. P. Gygi. 2016. Defining the consequences of genetic variation on a proteome-wide scale. *Nature.* **534:** 500–505.

28. Kettunen, J., T. Tukiainen, A. P. Sarin, A. Ortega-Alonso, E. Tikkanen, L. P. Lyytikainen, A. J. Kangas, P. Soininen, P. Wurtz, K. Silander, et al. 2012. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **44:** 269–276.

29. Muniandy, M., V. Velagapudi, A. Hakkarainen, J. Lundbom, N. Lundbom, A. Rissanen, J. Kaprio, K. H. Pietilainen, and M. Ollikainen. Plasma metabolites reveal distinct profiles associating with different metabolic risk factors in monozygotic twin pairs. *Int. J. Obes. (Lond.).* Epub ahead of print. June 15, 2018; doi:10.1038/s41366-018-0132-z.

30. Rhee, E. P., Q. Yang, B. Yu, X. Liu, S. Cheng, A. Deik, K. A. Pierce, K. Bullock, J. E. Ho, D. Levy, et al. 2016. An exome array study of the plasma metabolome. *Nat. Commun.* **7:** 12360.

31. Yet, I., C. Menni, S. Y. Shin, M. Mangino, N. Soranzo, J. Adamski, K. Suhre, T. D. Spector, G. Kastenmuller, and J. T. Bell. 2016. Genetic influences on metabolite levels: a comparison across metabolomic platforms. *PLoS One.* **11:** e0153672.

32. Vickers, K. C., B. T. Palmisano, B. M. Shoucri, R. D. Shamburek, and A. T. Remaley. 2011. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* **13:** 423–433.

33. Thomou, T., M. A. Mori, J. M. Dreyfuss, M. Konishi, M. Sakaguchi, C. Wolfrum, T. N. Rao, J. N. Winnay, R. Garcia-Martin, S. K. Grinspoon, et al. 2017. Adipose-derived circulating miRNAs regulate gene expression in other tissues. *Nature.* **542:** 450–455.

34. Arroyo, J. D., J. R. Chevillet, E. M. Kroh, I. K. Ruf, C. C. Pritchard, D. F. Gibson, P. S. Mitchell, C. F. Bennett, E. L. Pogosova-Agadjanyan, D. L. Stirewalt, et al. 2011. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. USA.* **108:** 5003–5008.

35. Kakugawa, S., P. F. Langton, M. Zebisch, S. Howell, T. H. Chang, Y. Liu, T. Feizi, G. Bineva, N. O'Reilly, A. P. Snijders, et al. 2015. Notum deacylates Wnt proteins to suppress signalling activity. *Nature.* **519:** 187–192.