

Causes of early-onset type 1 diabetes: toward data-driven environmental approaches

Pierre Bougnères and Alain-Jacques Valleron

A new study reveals distinctive metabolic changes that precede the development of type 1 diabetes (T1D), tossing a stone into the quiet waters of T1D immunology and genetics. The causes of these metabolic changes and their relationship to autoimmunity and β cell destruction are not yet known, but the identification of a metabolic phenotype linked to susceptibility to type 1 diabetes may help pave the way to a new era of investigation of T1D causality.

Autoimmunity is the predominant effector mechanism of T1D, but not its primary cause. Why the immune system becomes activated and attacks β cells in certain individuals is not clear. The causes of T1D include both intrinsic and extrinsic factors involving the patients' genotype, epigenotype, and environment.

That autoimmune processes drive the massive β cell destruction in certain forms of diabetes mellitus has been recognized for more than 30 yr. During the 1920s and 1960s, several pathologists reported "inflammation involving the islands of Langerhans" (insulinitis) in juvenile T1D. In the mid-1970s, autoantibodies against pancreatic islet cell antigens were identified in the sera of patients who had recently developed T1D (1), and a predisposing role for certain *HLA* genotypes was established (2, 3). However, despite the approximately 18,000 papers on T1D etiology that have followed these initial discoveries, the precise causes and initiating events that drive this disease in humans remain unclear.

In sibs of young T1D patients, autoantibodies against unidentified islet cell antigens were found to predict failure of insulin secretion and hyperglyce-

mia (4). The major β cell antigens recognized by autoantibodies were later identified as insulin, glutamic acid decarboxylase (GAD), and protein tyrosine phosphatase receptor type N (IA-2 proteins), and these antigens are now used for routine serological testing. Young children whose B cells produce significant amounts of these autoantibodies will usually, but not always, experience a subacute decline of β cell function and will soon become hyperglycemic (5). And although specific subsets of T lymphocytes are known to be key players in T1D, their precise place in the natural history of the disease remains to be elucidated.

On page 2975 of this issue, Orešič et al. present intriguing data that may provide hints about the etiology of this disease. The authors report characteristic changes in specific serum metabolites that precede the appearance of GAD- and insulin-specific autoantibodies and other metabolic changes that predict which children will eventually develop T1D (6).

Metabolic prelude to T1D

Orešič et al. compared serum metabolite profiles between children who eventually developed T1D and those who remained healthy and autoantibody-free (6). This analysis uncovered characteristic metabolic changes only in

the children who later developed T1D, including reduced serum succinate, phosphatidylcholine, and phospholipids, as well as decreased ketoleucine and elevated glutamic acid.

This sort of metabolomic approach to T1D natural history may be a pioneering example of new environmental data-driven approaches. It is not yet clear how higher levels of circulating glutamate, succinate, ketoleucine, or branched-chain amino acid might normally interfere with T1D initiation. It is also difficult to envision how these changes might have occurred. They could reflect an asymptomatic infection of the liver or muscle, a dietary event, or a metabolic disturbance in response to environmental cues.

Glutamate, which was elevated in children who went on to develop diabetes, is present in food, but presumably not in quantities capable of increasing serum levels to the extent reported (~32-fold above normal). Elevated serum glutamate in conjunction with diminished α -ketoglutarate could also result from increased muscle catabolism or impairment of the liver glutamate-dehydrogenase pathway and ureagenesis (7). It is also important to keep in mind that plasma glutamine and glutamate comprise only a small fraction of total intracellular glutamate pools and do not necessarily reflect the kinetics of whole-body fluxes (8). Humans who consume 10 g or more of monosodium glutamate have double the normal plasma glutamate levels and higher insulin concentrations (9). Although the total glutamate intake in Europe ranges from 5 to 12 g/d (10), glutamate is also endogenously synthesized with a turnover

P.B. is at Department of Pediatric Endocrinology and U561 INSERM, Saint Vincent de Paul Hospital, Paris V René Descartes University, 75014 Paris, France.

A.-J.V. is at Unité de Santé Publique and U707 INSERM, Saint Antoine Hospital, Paris VI Pierre et Marie Curie University, 75012 Paris, France.

CORRESPONDENCE

P.B.: bougneres@paris5.inserm.fr

of 48 g/d in adults. β cells are equipped with glutamate receptors and transporters and respond to glutamate with enhanced insulin secretion, suggesting a role for this metabolite in insulin regulation. However, the small overall mass of β cells makes them unlikely to be a significant contributor to increased serum glutamate (11). It is likely that an increased glutamate load in β cells increases the activity of GAD65, one of the major β cell antigens. Thus, one can hypothesize that transiently elevated glutamate and increased GAD activity might trigger or accelerate β cell damage through cytolytic or autoimmune processes. This process might be akin to autoimmunity of the thyroid, in which a higher iodine supply increases thyroperoxidase activity and the incidence of autoimmune thyroiditis (12).

Glutamate seems to be important to the microbiome composition, and some researchers speculate that *Escherichia coli* survive in the gut because of their GadABC and CadBA glutamate- and lysine-dependent acid-resistance systems (13). Glutamine and glutamate also have direct, but still unclear, effects on the immune system (14).

Increased levels of lysophosphatidylcholine (LPC) were detected by Orešič et al. in future T1D children at birth and during their first years (6). These lipid changes might result from events starting in utero, possibly linked to the mother's nutrition and metabolism. In addition, LPC, which is a bioactive byproduct of phospholipase A₂ (PLA₂) and a constituent of oxidized low-density lipoprotein, can influence the chemotaxis of leukocyte subpopulations during inflammation (15). And group VIA phospholipase A₂ (iPLA₂ β) participates in insulin secretion, as revealed by pharmacological or genetic manipulation of iPLA₂ β activity in β cells (16). For simplicity, in this article, we have restricted this discussion to glutamate and LPC variations in pre-T1D children, but other components of the metabolome might also be important for the genesis of the prediabetic state.

Epidemiological enigmas

The implications of the study by Orešič et al. are potentially important (6),

particularly given that T1D is on the rise in very young children, where the incidence has doubled in the last 20 yr (17). So far, no genetic or environmental factors have emerged to explain this pediatric phenomenon, but the fact that autoantibodies appear during the first 2 yr of life implicates early events (5). Emerging or rapidly evolving environmental changes may be to blame, but Ariadne's clue is missing.

This sort of metabolomic approach to T1D natural history may be a pioneering example of environmental data-driven approaches.

Another epidemiological enigma is the major difference in T1D prevalence across countries of the old continent. For example, Sardinia and Finland have the highest incidence of T1D worldwide (\sim 31/100,000 and 36/100,000, respectively), which is 5–8 times the incidence in other European countries. Both of these countries have distinctive genetic, cultural, and environmental characteristics, which thus far have defied the imaginations of T1D researchers.

Working hypotheses to explain the etiology of T1D are few and far between. One is the “hygiene hypothesis” (18), which suggests that fewer infections in modern day children results in decreased immune defenses and, as a result, increased incidence of childhood-onset diseases—asthma and T1D in particular (19). This may be an attractive hypothesis, but it has not yet been supported by scientific data. And notable examples of infections used to support the hygiene hypothesis often include prototypical infectious diseases, such as rheumatic fever, tuberculosis, mumps, measles, and hepatitis A (19). However, these diseases tend to occur later in life, and thus may not be the best reflection of pathogen exposure during early infancy. Indeed, the immune systems of children today are largely trained by frequent viral encounters in schools, urban environ-

ments, and through travel. Thus, it is unlikely that the hygiene hypothesis can solely explain the rapid rise of T1D cases in very young children. However, we may be able to find traces of relevant microbiological encounters in oropharyngeal or gut microbiomes by testing serological markers, or by quantifying the burden of childhood infection epidemics (20).

A shift toward hypothesis-generating research

Marcel Proust wrote that “discovery needs new eyes more than new lands.” T1D research certainly needs both. The new lands might include factors that affect a child's environment from periconception to disease onset, including environmentally sensitive epigenetics, microbiomics (21), metabolic events, and specific dietary changes. The new eyes of T1D research will likely include epigenetic epidemiology (22), serological markers of infections and metagenomics (23), metabolomics (24), and characterization of social ties (25). By combining several of these techniques and appropriate statistical analyses, it may be possible to start characterizing the T1D “environmentome.” We suggest that the development of whole-environment scans—data-driven investigation of a person's milieu using all available information on geographical location and personal history—will facilitate this process. Such a data-driven strategy will help us discover causal environmental factors and will extend classical hypothesis-driven epidemiology.

To facilitate this shift, medical researchers will have to develop new tools to examine an individual's personal geography and history and that will enable the individual's environment to be mapped. Some of these tools (cohort studies, computational mathematics, and statistics) will likely be inspired by those forged to explore infectious disease transmission or social networks (25). Epigenetics, microbiome and microbiology, metabolism and alimentation, and climatic environmental exposures are often interconnected (26). Seeking relationships between the geography of human and animal movements with

trajectories of infectious diseases, for example, will shed light on events at the population level. An eventual readout of an individual's predisposition to T1D should be derived from a combination of genetic and environmental factors. Indeed, certain epigenetic marks are dependent on specific food components (27), xenobiotics, pollutants, and metals. Such analyses will undoubtedly uncover unexpected and intriguing associations. For example, cold-chain development paralleled the Crohn's disease outbreak of the 20th century (28). The observation that old refrigerators are a major risk factor for disease may seem bizarre until one considers that bacteria identified in Crohn's disease were found to live at low temperatures.

Epigenetics, microbiome and microbiology, metabolism and nutrient intake, and climatic and environmental exposures are all dynamic. Each undergoes rapid adaptive changes during infancy and early childhood, thus contributing to plasticity at the organism level, including the developing immune system. If these time-dependent changes are stochastic, finding causative needles-in-the-haystack will be a great challenge. If the stochastic nature depends mostly on somatic and tissue-specific forces impinging on gene expression and protein activity, the task will be formidable.

By combining several of these techniques and appropriate statistical analyses, it may be possible to start characterizing the T1D "environmentome."

Studies might thus have to be longitudinal to enable the observation of relevant time-dependent changes. "Hit-and-run" encounters with environmental agents will not always leave detectable traces. Adequate control populations will have to be defined for each of the approaches. Another challenge is that all approaches will initially have to be blind (rather than "hypothesis-driven"), and thus will require

patient samples on an epidemiological scale. Findings gleaned from laboratory mice will not help at this early phase. Although many scientific minds object to such "fishing expeditions," these forays will be the inevitable first steps before hypotheses can be generated and tested in patients and experimental models.

After all, genome-wide association analyses of T1D genetics have become highly respected fishing expeditions across the hundreds of thousands of single-nucleotide polymorphisms that are now available to tag the individual genome. It took 20 yr to devise appropriate tools for investigating individual genome variability through genotyping techniques, genomic statistics, and genetic epidemiology, starting with either testing specific candidate genes selected by investigators or with low-powered microsatellite markers of the whole genome. We are now close to obtaining a full characterization of whole genomic and epigenomic variations. Similarly, it will take time to benefit fully from a systematic exploration of all available population and hospital databases providing historical and/or prospective information on the history of infections, drug use, and natural or industrial environmental exposures.

Genetics limits and new challenges

Genetic predisposition provides only a limited part of T1D causality, as 50–70% of monozygous twins are not concordant for the disease, and migrant children seem to adopt the T1D incidence of the country where they live rather than from where they originate. Genetics are not expected to contribute to the current increase of T1D incidence in young children, in whom *HLA* predisposition may even have diminished (29). Pure genetic analysis currently disregards the complexities of gene–environment interactions. However, genetic predisposition is only indicated at two levels: interaction with a hypothesized environmental factor (where, for example, the phenotype to be tested is a viral infection, not T1D) or development of T1D after interaction with the virus.

Even though the difficulties are monumental, the time may be ripe to encourage these approaches because they will enable research efforts to go from disease causes to disease mechanisms.

Future epidemiological investigations that attempt to dissect these tricky gene–environment interactions might benefit from using Mendelian randomization in which a functional genetic variant acts as a proxy for an environmental exposure (30). Currently, establishing causal relationships between environmental exposures and common diseases is beset with problems of unresolved confounding, reverse causation and selection bias that might result in spurious inferences. Mendelian randomization provides a means of overcoming these problems, because the inheritance of genetic variant is independent of (i.e., randomized in respect to) the inheritance of other traits, according to Mendel's law of independent assortment. Regarding the study by Orešič et al. (6), it is conceivable that genetic variants having sufficient influence on glutamate or lysophosphatidyl choline metabolism may help confirm the postulated implication of these metabolites in T1D.

Few researchers have started to match genetic diversity with environmental variety. *HLA* variants (which have their own geographical distribution) could influence T1D susceptibility, predominantly in specific clustered environments (31). Regional population differences, which presumably include lifestyle and geographic and biotic factors, have been shown to modulate genome-wide gene expression to the same extent as genetic divergence (32). DNA changes associated with T1D may not lead directly to disease, but instead may act on intermediate molecular phenotypes that in turn induce modifications in high-order disease traits. Recently, gene-expression genetics ("expression mapping") in human populations has

started to provide a more objective view of the contribution of genetic variation to variations in gene expression—thousands of traits have been found to be under the control of well-defined genetic loci. The genetics of gene expression have been studied in human lymphoblastoid cell lines and human liver, where T1D-associated single-nucleotide polymorphisms in genome-wide scans have also been associated with liver gene-expression traits (33). Not only will this approach enable us to understand the biology underlying the statistical association of a genomic variant with a complex disease, it will ultimately help determine which parts of the gene-expression variation are attributable to genetic or nongenetic factors. The same kind of gene-expression studies can be undertaken to test the effect of environmental exposures in relevant populations. However, the problem persists of the tissue where gene expression is expected to be relevant to the disease, such as in lymphocytes for asthma or T1D (34).

The challenges of new lands

Metabolomics, metagenomics, microbiology, epigenetic epidemiology, and gene-expression studies will face the same high-throughput challenges that whole-genome strategies did. The utilization of accurate techniques for characterizing individual markers, appropriate case-control comparisons, the sophistication of computational mathematics, and the precautions required to accurately interpret the data will be required to overcome the blindness of the experimental design and the complexity and heterogeneity of T1D causality. Even though the difficulties are monumental, the time may be ripe to encourage these approaches because they will enable research efforts to go from disease causes to disease mechanisms. This strategy should offset the reliance on animal models, not only because of their own limitations but also because these models drive ideas and studies in only one direction: from disease mechanisms to causes. Incentives to use large fishing expeditions in man will be high because, in addition to contributing to our understanding of T1D complexity,

these hypothesis-generating approaches should provide significant clues to the prediction and prevention of T1D at an individual or a population level.

We thank J. Jacobson for editorial help and NovoNordisk France for supporting this work.

REFERENCES

- Bottazzo, G.F., A. Florin-Christensen, and D. Doniach. 1974. Islet-cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. *Lancet*. ii:1280–1283.
- Nerup, J., P. Platz, O.O. Andersen, M. Christy, J. Lyngsoe, J.E. Poulsen, L.P. Ryder, L.S. Nielsen, M. Thomsen, and A. Svegaard. 1974. HL-A antigens and diabetes mellitus. *Lancet*. ii:864–866.
- Benacerraf, B., and H.O. McDevitt. 1972. Histocompatibility-linked immune response genes. *Science*. 175:273–279.
- Srikanta, S., O.P. Ganda, G.S. Eisenbarth, and J.S. Soeldner. 1983. Islet-cell antibodies and beta-cell function in monozygotic triplets and twins initially discordant for Type I diabetes mellitus. *N. Engl. J. Med.* 308:322–325.
- Ziegler, A.G., M. Hummel, M. Schenker, and E. Bonifacio. 1999. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes*. 48:460–468.
- Orešič, M., S. Simell, M. Sysi-Aho, K. Näntö-Salonen, T. Seppänen-Laakso, V. Parikka, M. Katajamaa, A. Hekkala, I. Mattila, P. Keskinen, et al. 2008. Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *J. Exp. Med.* 205:2975–2984.
- Hack, V., O. Stütz, R. Kinscherf, M. Schykowski, M. Kellerer, E. Holm, and W. Dröge. 1996. Elevated venous glutamate levels in (pre)catabolic conditions result at least partly from a decreased glutamate transport activity. *J. Mol. Med.* 74:337–343.
- Darmaun, D., D. Matthews, and D.M. Bier. 1986. Glutamine and glutamate kinetics in humans. *Am. J. Physiol. Endocrinol. Metab.* 251: E117–E126.
- Mourtzakis, M., and T.E. Graham. 2002. Glutamate ingestion and its effects at rest and during exercise in humans. *J. Appl. Physiol.* 93: 1251–1259.
- Beyreuther, K., H.K. Biesalski, J.D. Fernstrom, P. Grimm, W.P. Hammes, U. Heinemann, O. Kempster, P. Stehle, H. Steinhart, and R. Walker. 2007. Consensus meeting: monosodium glutamate, an update. *Eur. J. Clin. Nutr.* 61:304–313.
- Kelly, A., C. Li, Z. Gao, C.A. Stanley, and F.M. Matchinsky. 2002. Glutaminolysis and insulin secretion. *Diabetes*. 51:S421–S426.
- Teng, W., Z. Shan, X. Teng, H. Guan, Y. Li, D. Teng, Y. Jin, X. Yu, C. Fan, W. Chong, et al. 2006. Effect of iodine intake on thyroid diseases in China? *N. Engl. J. Med.* 354:2783–2793.
- Moreau, P.L. 2007. The lysine decarboxylase CadA protects *Escherichia coli* starved of phosphate against fermentation acids. *J. Bacteriol.* 189:2249–2261.
- Newsholme, P. 2001. Why is glutamine metabolism important to cells of the immune system in health, postinjury, surgery or infection. *J. Nutr.* 131:2515S–2522S.
- Murugesan, G., M.R. Sandhya Rani, C.E. Gerber, C. Mukhopadhyay, R.M. Ransohoff, G.M. Chisolm, and K.J. Kottke-Marchant. 2003. Lysophosphatidylcholine regulates human microvascular endothelial cell expression of chemokines. *J. Mol. Cell. Cardiol.* 35: 1375–1384.
- Bao, S., H. Song, M. Wohltmann, S. Ramanadham, W. Jin, A. Bohrer, and J. Turk. 2006. Insulin secretory responses and phospholipid composition of pancreatic islets from mice that do not express Group VIA phospholipase A2 and effects of metabolic stress on glucose homeostasis. *J. Biol. Chem.* 281:20958–20973.
- Green, A., and C.C. Patterson. 2001. Trends in the incidence of childhood-onset diabetes in Europe 1989–1998. *Diabetologia*. 44:B3–B8.
- Strachan, D.P. 1989. Hay fever, hygiene, and household size. *BMJ*. 299:1259–1260.
- Bach, J.F. 2002. The effects of infections on susceptibility to autoimmune and allergic diseases. *N. Engl. J. Med.* 347:911–920.
- Flahault, A., T. Blanchon, Y. Dorléans, L. Toubiana, J.F. Vibert, and A.J. Valleron. 2006. Virtual surveillance of communicable diseases: a 20-year experience in France. *Stat. Methods Med. Res.* 15:413–421.
- Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, and D.A. Relman. 2005. Diversity of the human intestinal microbial flora. *Science*. 308:1635–1638.
- Bjornsson, H.T., M.D. Fallin, and A.P. Feinberg. 2004. An integrated epigenetic and genetic approach to common human disease. *Trends Genet.* 20:350–358.
- Gill, S.R., M. Pop, R.T. Deboy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, D.A. Relman, C.M. Fraser-Liggett, and K. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. *Science*. 312:1355–1359.
- Holmes, E., I.D. Wilson, and J.K. Nicholson. 2008. Metabolic phenotyping in health and disease. *Cell*. 134:714–717.
- Christakis, N.A., and J.H. Fowler. 2007. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* 357:370–379.
- Turnbaugh, P.J., and J.I. Gordon. 2008. An invitation to the marriage of metagenomics and metabolomics. *Cell*. 134:708–713.
- Van den Veyver, I.B. 2002. Genetic effects of methylation diets. *Annu. Rev. Nutr.* 22: 255–282.
- Hugot, J.P., C. Alberti, D. Berrebi, E. Bingen, and J.P. Cézard. 2003. Crohn's disease: the cold chain hypothesis. *Lancet*. 362:2012–2015.
- Gillespie, K.M., S.C. Bain, A.H. Barnett, P.J. Bingley, M.R. Christie, G.V. Gill, and E.A.

- Gale. 2004. The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. *Lancet*. 364:1699–1700.
30. Ebrahim, S., and G. Davey Smith. 2008. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum. Genet.* 123:15–33.
31. Kukko, M., S.M. Virtanen, A. Toivonen, S. Simell, S. Korhonen, J. Ilonen, O. Simel, and M. Knip. 2004. Geographical variation in risk HLA-DQB1 genotypes for type 1 diabetes and signs of beta-cell autoimmunity in a high-incidence country. *Diabetes Care*. 27:676–681.
32. Idaghdour, Y., J.D. Storey, S.J. Jadallah, and G. Gibson. 2008. A genome-wide gene expression signature of environmental geography in leukocytes of Moroccan Amazighs. *PLoS Genet.* 4:e1000052.
33. Schadt, E.E., C. Molony, E. Chudin, K. Hao, X. Yang, P.Y. Lum, A. Kasarskis, B. Zhang, S. Wang, C. Suver, et al. 2008. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 6:e107.
34. Moffatt, M.F., M. Kabisch, L. Liang, A.L. Dixon, D. Strachan, S. Heath, M. Depner, A. von Berg, A. Bufe, E. Rietschel, et al. 2007. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature*. 448:470–473.