Published in final edited form as:

Med Hypotheses. 2010 February ; 74(2): 289. doi:10.1016/j.mehy.2009.09.009.

Accounting for Chance in the Calculus of Autoimmune Disease

Daniel J. Moore¹

¹Department of Pediatrics, Ian Burr Division of Endocrinology and Diabetes, Vanderbilt Children's Hospital, Nashville, TN 37232, USA

Abstract

Discussions around the etiology of autoimmune disease routinely focus on the interplay between genes and the environment. In turn, efforts to ameliorate these diseases seek to modify genetic and environmental factors. However, there may be a third element that also accounts for the progression of autoimmunity. This article explores the role of chance, exemplified by the stochastic process of immune repertoire generation, in the evolution of autoimmunity. The presented modeling studies and concepts suggest that chance plays as significant a role as genes or environment. This hypothesis implies that a full understanding of the role of genes and environment will also require investigators to account for stochastic processes in building comprehensive disease models.

Keywords

Autoimmunity; T cell; B cell; Diabetes; Twin; Environment; TCR; Stochastic Process; Genetics; T cell repertoire

Introduction

A common theme in the discussion of autoimmune diabetes and most other autoimmune diseases is the relative contribution of genes and the environment. While important contributions from both are well appreciated, their relative contributions continue to be debated. Studies of identical twins have provided a significant body of evidence that fuels this discussion. In several series of twins, the concordance of Type 1 (autoimmune) diabetes has been reported as high as 60% [1-4] and may be even higher if the discordant twin is followed for an entire lifetime. This finding has been interpreted to indicate that up to 60% of the risk for diabetes development is genetically conferred with the remaining 40% being conferred by as yet unidentified environmental agents.

These data have led to an extensive search to find these environmental agents and also to many efforts to modify the environment in order to prevent disease [5-11]. However, the strict division between genes and environment is dissatisfying. In particular, the environment shared by twins is often more similar than even between other siblings. Moreover, these interpretations do not account for a fundamental property of the immune system—the random generation of receptors

© 2009 Elsevier Ltd. All rights reserved.

Corresponding author: Daniel J. Moore, Vanderbilt Children's Hospital, Department of Pediatrics, Ian Burr Division of Endocrinology and Diabetes, 2200 Children's Way, 11136 Doctors' Office Tower, Nashville, TN, 37232-9170, USA, daniel.moore@vanderbilt.edu.

Conflict of Interest Statement The author has no conflicts of interest to disclose

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

by genetic recombination. Therefore, the hypothesis explored in this manuscript is that the stochastic process of immune selection represents a quantifiable factor that acts to determine autoimmune disease development.

Evaluation of the hypothesis

Randomly generated T and B lymphocytes and their interactions play a central role in the pathogenesis of Type 1 diabetes[12-20]. In all individuals, these cells are subjected to highly efficient developmental processes that attempt to regulate or eliminate autoreactive clones. We questioned what the impact of different frequencies of “escape from selection” (failed tolerance) might have on the development of an immune repertoire. This analysis, presented below, suggests that at biologically plausible frequencies of escape, a pair of twins will not have both generated an autoreactive repertoire during childhood. As the peak age of diabetes onset is age 12, persons with high genetic susceptibility to autoimmunity may avoid disease not because of differences in the environment but simply because of the basic properties of the immune system.

Probability of generating an islet-reactive specificity

We began by determining an appropriate range for the potential of an autoreactive cell to escape to the periphery during a single day (Pescape) for the purposes of modeling the repertoire over time. The chance of such a cell escaping to the periphery itself includes at least two factors, the chance of generating such a cell (Pgen) and the subsequent chance that negative selection fails to eliminate it (Pfail). For our initial analysis, we will focus on the final outcome, the chance of escape to the periphery, and will ignore initially the relative contributions of generation and selection. We will address these factors in a later section since appreciating the possible range for these factors is crucial in assessing the biologic plausibility of the model. For the purposes of this analysis, we will focus on the generation of T lymphocytes when we discuss specifics of immune generation such as numbers of cells generated per day and the diversity of the repertoire. Similar logic would apply to B lymphocyte selection.

First, we will establish the probability of generating a single autoreactive clone as a function of time and will establish a range of Pescape that will include 100% chance of generating these clones down to very minuscule probabilities. Based on a range of possibilities for Pescape, we have applied the following formula to determine the long-term prospects of generating an autoreactive clone.

$$\text{Chance of one Autoreactive Clone in Periphery} = 1 - (1 - P_{\text{escape}})^{(365 \times \text{Years})}$$

As shown in Figure 1, the chances of generating an autoreactive clone during childhood (the first 10-13 years of life) change rapidly over a relatively narrow range of escape probabilities. At a Pescape of 0.01, escape is almost immediate. At a chance of 0.001, there is a 97% chance of generating an autoreactive clone by age 10. At ten-fold less, there is only a 50% chance of generating such a clone in the first 19 years of life. At a chance of escape of 1 in a million per day, there is a less than 1% chance of an autoreactive clone escaping to the periphery in the first 19 years of life.

Probability of generating a quorum of islet-reactive cells

Allowing just one autoreactive cell into the periphery is unlikely to convey much risk for the development of overt autoimmune disease. In fact, it is well established that contributions from the randomly generated CD4 and CD8 T cell and B cell compartments are all needed for the development of autoimmunity[12-15,17,21,22]. In addition, predictive models of diabetes

development in humans suggest that the risk for long-term diabetes development increases dramatically once at least two autoantibodies are present in the peripheral serum[23-25]. As these two autoantibodies would come from two distinct B cells, at least two autoreactive B cells must be produced. B lymphocytes act in the development of diabetes by presenting antigen via the MHC complex to islet-reactive T cells[13]. Thus, these B cells likely need specific partners to cause disease. Overall, a bare minimum of three to six independently generated autoreactive cells will be needed for the development of Type 1 diabetes. We have thus considered the probability of generating more than one autoreactive cell as a function of time. As shown in Figure 2, we have determined the chances of generating either three or six such clones at the same range of Pescape values used in Figure 1. We have also shown the chance of generating one to ten autoreactive clones by the age of 13, a time just after the peak incidence of Type 1 diabetes (Figure 2c). Figure 2a/b demonstrate that the chance of generating a quorum of autoreactive clones increases dramatically over a relatively narrow (ten-fold) range of Pescape values. At an efficiency of 1:10,000, there is minimal (<5%) chance of generating such a quorum in childhood. With an increase to 1:1,000, that chance leaps to 97% by age 13.

Assessing the relative contributions of production and selection to the chance of generating an autoreactive clone

We have established a range of values for Pescape that result in a chance of autoreactive cell generation from near certainty to only a slight possibility. By analyzing the role of a quorum of autoreactive cells, the range of Pescape is further constrained. The data suggest that at Pescape of 1:10,000 protection is near complete. Conversely, at a likelihood of escape to the periphery of 1:1,000, the production of an autoimmune repertoire is nearly assured. To determine whether the proposed escape probabilities are biologically feasible, we will consider the relative contribution of two important determining factors: the chance that a clone deriving from the bone marrow is autoreactive and the chance that negative selection fails to eliminate the clone. Each day the immune system generates numerous new cells from the combination of random generation of T cell receptors (TCR) from recombination of TCR alpha and beta chains in the bone marrow. Generation of these cells is then followed by expansion through proliferation. Most estimates suggest that 10^6 to 10^7 unique clones are formed each day and then expanded. While we do not know exactly what proportion of the possible TCRs would be autoreactive, it is likely that at least some autoreactive cells are generated every day. A range of Pgen from 0.000001 (implying one in a million recombined TCRs are autoreactive) to .01 (implying one in one hundred recombined TCRs are autoreactive) would allow for between 1 and 10,000 autoreactive cells to be generated each day from a starting pool of one million newly formed cells.

Following the generative process, these cells are subject to selection. The selection process is highly efficient and up to 98% of all generated T cells fail this process and are eliminated. This process is inherently biased such that non-autoreactive cells are more likely to survive to the periphery. However, the degree of efficiency may vary based on genetics and may be an underpinning of autoimmunity. The selection efficiency can be related to the probability of escape based on the number of cells that challenge the selection machinery each day as follows:

$$1P_{\text{escape}} = (1P_{\text{fail}})^{(\text{autoreactive clones per day})}$$

$$1P_{\text{escape}} = (1P_{\text{fail}})^{((\text{total unique clones produced per day}) \times (P_{\text{gen}}))}$$

Based on the analysis in the previous sections, a likelihood of escape near 1:10,000 affords good protection against the development of a quorum of autoreactive specificities. Therefore, we have considered how the two contributing factors must vary to produce Pescape in this neighborhood (Figure 3). For this purpose, we have assumed that 1 million unique clones are produced daily; application of the Pgen frequency then determines the number of autoreactive

cells arising per day. The presented analysis allows from 1 to 10,000 autoreactive cells to emerge each day to test the selection machinery. Because the range of *Pescape* varies from very small numbers to 1, we have plotted the outcomes in figure 3 on a log scale. We have set up the scale such that values under 0 represent *Pescape* $<1:10,000$ and thus a low likelihood of producing an autoimmune repertoire. Conversely, values over 0 represent *Pescape* $>1:10,000$ and ever increasing chances of autoimmune repertoire production. In order to obtain an acceptable level of *Pescape* based on the data presented in Figures 1 and 2, up to 100 autoreactive clones can be produced per day without compromising immune tolerance. At this frequency of repertoire challenge, the selection machinery must fail to eliminate the autoreactive clone only 1 in a million times to maintain *Pescape* within the safe range.

Probability of Twin Concordance

Finally, we have estimated the chances that two individuals operating under the same efficiency of autoreactive clone selection would both possess an autoreactive quorum during childhood. Clinically, this parallels our earlier question of how likely it might be that identical twins would be equally susceptible to disease induction if facing the same environmental trigger. We have graphed the possibility that both twins would possess a quorum of either three or six autoreactive clones as a function of age as shown in Figure 4. Similar to Figure 2, at *Pescape* below $1:10,000$ there is almost no chance of concordance. However, at a level of escape of $1:1,000$, there is about an 80% chance of both twins having a susceptible repertoire at age 9. At other levels between $1:1,000$ and $1:10,000$ there is a range of potential concordances with many values around 50%. Thus, even if both individuals encounter the same environmental trigger, there is less than a 100% chance that both will be able to progress to disease. Discordance in disease progression may therefore relate as much to the state of the immune system as to the presence of environmental triggers.

Discussion and Implications of the Hypothesis

There are clearly important contributions to the development of autoimmune diabetes from both genes and the environment. However, we have considered a third contributing factor here, the role of the basic fundamental of immune system function which produces T and B cells in a random fashion. We have considered the long-term outcome of the efficiency by which these cells are subjected to elimination on the formation of the immune repertoire and the risk for autoimmune disease.

A number of interesting features emerge from this consideration. When one imposes the additional constraint of needing a quorum of cells for the production of overt autoimmunity, a daily risk of autoreactive cell escape of $1:10,000$ provides near complete protection against the development of autoimmunity in the first ten years of life. A requirement of six such cells reduces the chances to almost nil. Thus, there is little incentive for the immune system to exhibit more stringent measures of selection. Interestingly, at a level of selection of $1:10,000$, there is a very significant chance of generating at least one autoreactive cell by age 19 (near 50%). This aspect of the data is in keeping with the general clinical experience that a large proportion of normal individuals will have at least one autoantibody, such as anti-DNA or rheumatoid factor. Further reductions in this value reduce the chance of observing autoimmune cells in the absence of autoimmune disease to levels that are not in keeping with this clinical experience.

It is also noteworthy how rapidly the risk for developing a quorum of autoimmune cells changes over a ten-fold range of selection stringency. Thus, small perturbations in selection may convey significant risk for autoimmune disease. This modeling is also in keeping with the clinical experience in which a large number of genes contribute in apparently minor ways to the development of disease[26-28]. This model suggests that only a slight change would significantly alter the risk profile over a short period of time. This change could modify the

production of autoreactive clones from the bone marrow or could reduce the efficiency of the selecting machinery that eliminates these clones. Based on the analysis presented in Figure 3, 100 autoreactive cells can be generated each day and a safe immune system maintained provided that only 1 in a million autoreactive cells escapes elimination. If the number of autoreactive cells that challenge the selection machinery is increased to just 600, the chance of generating an autoreactive quorum rises to over 70%. Clearly, only small changes in the behavior of the immune system are needed to substantially modify an individual's propensity to autoimmune disease.

This analysis was motivated in part to understand whether the stochastic generation of the immune system could account for the discordance seen in twins. We have addressed this by considering whether a set of twins would both have a susceptible immune repertoire at the time of environmental challenge. Our analysis suggests that there a range of selection efficiencies over which the twins would not both be susceptible during the peak age of diabetes onset. Thus, even if they experienced the same environmental stimulant, it is likely that both would not progress to diabetes.

As an important corollary of this hypothesis, we would advise that it is not possible to examine environmental triggers without knowing the state of the immune system including whether the individual under consideration has generated a quorum of autoreactive cells. The read-out of whether or not diabetes develops is not adequate for identifying diabetes risk factors from the environment and without this additional step, previous analysis of these risk factors are likely to have both over and under-estimated the risk of numerous environmental agents. This problem in data analysis may also account in part for the inability to consistently replicate these kinds of data between centers and over time.

At present it is not possible to fully test this model empirically. However, with technological advances such as MHC tetramers, we may soon know the range and identity of autoreactive specificities in some disease states. Thus, it may become possible to define a quorum. These data also suggest that an unaffected twin may be highly likely to develop some degree of autoimmunity over a lifetime. Long-term follow-up will test this corollary.

Acknowledgments

DJM is supported by NRSA training award 5 F32 DK083161-02 and by a grant from the Endocrine Fellows' Foundation.

References

1. Burch PR. Diabetes mellitus: Concordance in monozygotic twins. *Lancet* 1976;2:632. [PubMed: 61369]
2. Redondo MJ, Rewers M, Yu L, et al. Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study. *Bmj* 1999;318:698–702. [PubMed: 10074012]
3. Kyvik KO, Green A, Beck-Nielsen H. Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. *Bmj* 1995;311:913–917. [PubMed: 7580548]
4. Johnston C, Millward BA, Hoskins P, Leslie RD, Bottazzo GF, Pyke DA. Islet-cell antibodies as predictors of the later development of type 1 (insulin-dependent) diabetes. A study in identical twins. *Diabetologia* 1989;32:382–386. [PubMed: 2668086]
5. Akerblom HK, Knip M. Putative environmental factors in Type 1 diabetes. *Diabetes Metab Rev* 1998;14:31–67. [PubMed: 9605629]
6. Dahlquist G. Non-genetic risk determinants of type 1 diabetes. *Diabetes Metab* 1994;20:251–257. [PubMed: 8001712]

7. Karvonen M, Jantti V, Muntoni S, Stabilini M, Stabilini L, Tuomilehto J. Comparison of the seasonal pattern in the clinical onset of IDDM in Finland and Sardinia. *Diabetes Care* 1998;21:1101–1109. erratum appears in *Diabetes Care* 1998 Oct;21(10):1784. [PubMed: 9653603]
8. McKinney PA. Seasonality of birth in patients with childhood Type I diabetes in 19 European regions. *Diabetologia* 2001;44:B67–74. [PubMed: 11724420]
9. Kaila B, Taback SP. The effect of day care exposure on the risk of developing type 1 diabetes: a meta-analysis of case-control studies. *Diabetes Care* 2001;24:1353–1358. [PubMed: 11473069]
10. Robles DT, Eisenbarth GS. Type 1a diabetes induced by infection and immunization. *J Autoimmun* 2001;16:355–362. [PubMed: 11334504]
11. Dahlquist G, Mustonen L. Analysis of 20 years of prospective registration of childhood onset diabetes time trends and birth cohort effects. Swedish Childhood Diabetes Study Group. *Acta Paediatr* 2000;89:1231–1237. [PubMed: 11083381]
12. Serreze DV, Chapman HD, Varnum DS, et al. B lymphocytes are essential for the initiation of T cell-mediated autoimmune diabetes: analysis of a new “speed congenic” stock of NOD.Ig mu null mice. *J Exp Med* 1996;184:2049–2053. [PubMed: 8920894]
13. Noorchashm H, Lieu YK, Noorchashm N, et al. I-Ag7-mediated antigen presentation by B lymphocytes is critical in overcoming a checkpoint in T cell tolerance to islet beta cells of nonobese diabetic mice. *J Immunol* 1999;163:743–750. [PubMed: 10395666]
14. Yang M, Charlton B, Gautam AM. Development of insulinitis and diabetes in B cell-deficient NOD mice. *J Autoimmun* 1997;10:257–260. [PubMed: 9218752]
15. Noorchashm H, Moore DJ, Noto LE, et al. Impaired CD4 T cell activation due to reliance upon B cell-mediated costimulation in nonobese diabetic (NOD) mice. *Journal of Immunology* 2000;165:4685–4696.
16. Haskins K, Portas M, Bradley B, Wegmann D, Lafferty K. T-lymphocyte clone specific for pancreatic islet antigen. *Diabetes* 1988;37:1444–1448. [PubMed: 2458291]
17. Bradley BJ, Wang YY, Lafferty KJ, Haskins K. In vivo activity of an islet-reactive T-cell clone. *J Autoimmun* 1990;3:449–456. [PubMed: 1977404]
18. Zipris D, Lazarus AH, Crow AR, Hadzija M, Delovitch TL. Defective thymic T cell activation by concanavalin A and anti-CD3 in autoimmune nonobese diabetic mice. Evidence for thymic T cell anergy that correlates with the onset of insulinitis. *J Immunol* 1991;146:3763–3771. [PubMed: 1827815]
19. Miyazaki A, Hanafusa T, Yamada K, et al. Predominance of T lymphocytes in pancreatic islets and spleen of pre-diabetic non-obese diabetic (NOD) mice: a longitudinal study. *Clinical & Experimental Immunology* 1985;60:622–630. [PubMed: 3160515]
20. Hanafusa T, Sugihara S, Fujino-Kurihara H, et al. Induction of insulinitis by adoptive transfer with L3T4+Lyt2- T-lymphocytes in T-lymphocyte-depleted NOD mice. *Diabetes* 1988;37:204–208. [PubMed: 3292315]
21. Wong FS, Visintin I, Wen L, Flavell RA, Janeway CA Jr. CD8 T cell clones from young nonobese diabetic (NOD) islets can transfer rapid onset of diabetes in NOD mice in the absence of CD4 cells. *J Exp Med* 1996;183:67–76. [PubMed: 8551245]
22. Ou D, Jonsen LA, Metzger DL, Tingle AJ. CD4+ and CD8+ T-cell clones from congenital rubella syndrome patients with IDDM recognize overlapping GAD65 protein epitopes. Implications for HLA class I and II allelic linkage to disease susceptibility. *Hum Immunol* 1999;60:652–664. [PubMed: 10439311]
23. Verge CF, Gianani R, Kawasaki E, et al. Number of autoantibodies (against insulin, GAD or ICA512/IA2) rather than particular autoantibody specificities determines risk of type I diabetes. *J Autoimmun* 1996;9:379–383. [PubMed: 8816974]
24. Verge CF, Gianani R, Kawasaki E, et al. Prediction of type I diabetes in first-degree relatives using a combination of insulin, GAD, and ICA512bdc/IA-2 autoantibodies. *Diabetes* 1996;45:926–933. [PubMed: 8666144]
25. Barker JM, Barriga KJ, Yu L, et al. Prediction of autoantibody positivity and progression to type 1 diabetes: Diabetes Autoimmunity Study in the Young (DAISY). *J Clin Endocrinol Metab* 2004;89:3896–3902. [PubMed: 15292324]

26. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;39:857–864. [PubMed: 17554260]
27. Lowe CE, Cooper JD, Brusko T, et al. Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes. *Nat Genet* 2007;39:1074–1082. [PubMed: 17676041]
28. Hulbert EM, Smink LJ, Adlem EC, et al. T1DBase: integration and presentation of complex data for type 1 diabetes research. *Nucleic Acids Res* 2007;35:D742–746. [PubMed: 17169983]

Figure 1. Chance of generating one autoreactive clone

The graph displays the likelihood that an autoreactive clone will appear in the periphery in each of the first nineteen years of life over a range of daily probabilities for an autoreactive clone to be generated and escape to the periphery (Pescape). Each colored line represents a different Pescape ranging from 1:1,000,000 to 1:100.

Figure 2. Generating a quorum of autoreactive cells

The same analytical methods as in figure 1 are applied here, but are used to determine the chance of generating either three (A) or six (B) autoreactive clones as a function of time. Panel C illustrates the chance of generating up to ten autoreactive clones by age thirteen as a function of varying Pescape values.

Figure 3. Role of autoreactive cell generation and selection efficiency in determining overall production of autoreactive cells

The effect on Pescape of challenging the selection machinery by a range of potential autoreactive clones per day is plotted. The chance that selection will fail ranges from one in a million to one in ten thousand. The plot is a logarithmic display of the chance of escape. As determined from Figures 1 and 2, a Pescape of 1:10,000 represents good protection from autoimmunity. Therefore, we have plotted $\log (\text{Pescape} \times 10,000)$ so that values above 0 represent increasing risk of autoimmunity (Pescape greater than 1:10,000).

Figure 4. Concordance of Twin Repertoires at Multiple Levels of Clone Selection

The graph demonstrates the potential over time that a pair of twins will have both generated a quorum of autoreactive cells defined as either three (A) or six (B) such cells.