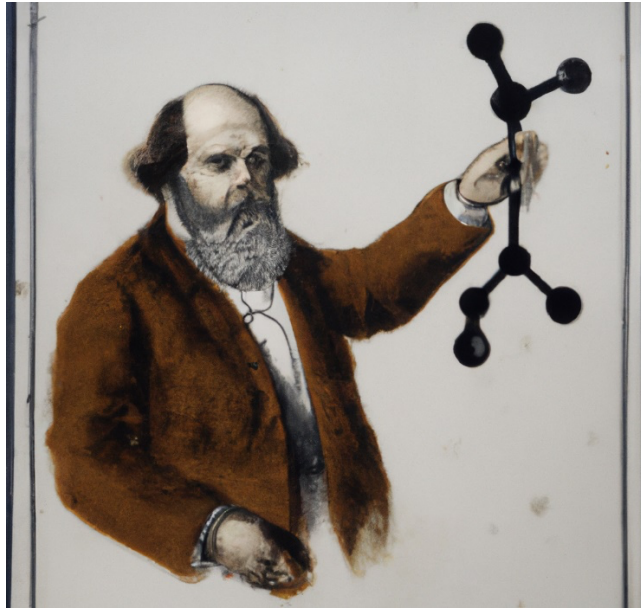


## On the Origins of Antibodies: Natural and Man-made



"As natural selection acts by life and death, by the survival or extinction of the less well-fitted individuals, and by the multiplication of the better fitted or adapted individuals, it necessarily follows that every slight modification of structure or instinct, which tends to enable an organism to compete better with its fellows, or to cope better with external conditions, will be preserved." [1]

Charles Darwin's groundbreaking theory of evolution fundamentally altered our understanding of the natural world. However, his profound insights predated the era of molecular biology, during which significant progress was made in elucidating the mechanisms by which genetic information is transmitted across generations via DNA, which serves as a blueprint encoding the proteins that comprise each species. It is worth noting that Darwin himself posited a mechanism of inheritance and, ironically, the term he coined for this mechanism – "gemmule" – was instrumental in the eventual coining of the term "gene" [2]. While Darwin's theory was initially conceived to account for the evolution of plants and animals, one cannot help but wonder what the great naturalist would have thought had he known how applicable this very same theory would be to the evolution of sub-cellular macromolecules. In fact, by simply substituting the terms "individual" and "organism" with "protein" in the quote above, it remains remarkably relevant to our understanding of protein evolution. This is particularly evident in the context of natural and synthetic antibody generation, which I will expound upon in the following paragraphs.

In "The Origin of Species" [1], Charles Darwin astutely observed that "As the number of individuals in any given species increases, so will the tendency to variability increase, and the chance of any new variety or deviation in structure arising will be greater." This insightful statement was further expounded upon in "The Variation of Animals and Plants Under Domestication," [3] where Darwin suggested that large groups of individuals provided a better chance for the emergence of favorable variations compared to smaller groups. Similarly, in "The Descent of Man," [4] he noted that "The larger the number of individuals, the better will be the chance for the appearance of favorable variations." These quotes serve to highlight Darwin's recognition of the importance of favorable variations in driving natural selection and the increased likelihood of such variations in larger populations.

The immune system, in an attempt to find this favorable variation in large numbers, generates a vast and diverse pool of B and T cell receptors to detect, neutralize, and eliminate invading foreign threats. The primary mechanism for generating diversity in B-cell receptors lies in the genome, with a multitude of variable (V), diversity (D), and joining (J) gene segments in the human genome for the immunoglobulin heavy chain and numerous V and J gene segments for the kappa and lambda light chains [5]. These genes rearrange in a nearly random fashion, creating the initial layer of diversity (combinatorial diversity) within the B-cell pool, including some V(DD)J recombined B-cells with larger CDRH3 loops [5,6]. Junctional diversity, the addition or deletion of bases where these genes meet, leads to another layer of variation that impacts the length of the final B-cell receptor's amino acid sequence, especially that of CDRH3 [7]. The differential pairing of different heavy and light chains, adds another layer of repertoire diversity. Finally, somatic hypermutation, the random introduction of point mutations into immunoglobulin variable regions by activation-induced cytidine deaminase, adds the last layer of diversity after B-cell receptor activation [8]. Class-switching further alters the functional properties of secreted antibodies, enabling antibodies to guide different functions in an immune response. These mechanisms result in a vast B-cell repertoire with as many as  $10^{18}$  members based on theoretical combinatorial calculations [9], a proverbial oil field if we refer to variation as the fuel of natural selection. Scientists attempting to emulate nature's large repertoires mimic this diversity in yeast, phage, or mammalian cell libraries. Antibody engineers worldwide strive to adhere to Darwin's principles, in an effort to increase their chances of yielding antigen binders, by aiming to create these artificial immune repertoires to be as large and as complex as possible- though they are still somewhat primitive in comparison to natural repertoires.

Returning to Darwin's opening quote, the survival and extinction of B-cell receptors is a fundamental aspect of natural repertoire development. Just as Darwin illustrated how Finches with smaller beaks were selected against and went extinct on Galapagos islands that offered larger seeds as a food source, B-cell receptors that are unfavorable to human survival are selected against through negative selection. In the case of receptor editing, B-cells with non-functional B-cell receptors due to frameshift mutations caused by improper joining are given another opportunity at life [10,11]. The B-cells are prompted to recombine their genes again to generate another functional B-cell receptor. If a B-cell produces a receptor that is highly reactive to self-antigens, apoptotic signals command the B-cell to undergo cell death, eradicating them from the repertoire [12].

The mechanism of B-cell expansion is akin to Darwin's explanation of variants that make a species better adapted to their environment, which outcompete and ultimately outnumber those that are less well-adapted. B-cells that bind antigens internalize the B-cell receptor/antigen complex, after which the antigen is processed and restricted antigen peptides are presented to T helper cells via MHC class II. The T-helper cells signal for the B-cell to proliferate and differentiate into antibody-secreting plasma cells. This process is continuous, with the best binding B-cell receiving more signals, and somatic hypermutation ensuring that more variation can be selected from. If somatic hypermutation generates a better binder, it receives more signal than its predecessor, but if it has slightly lower affinity, it receives less. Ultimately, the highest affinity binders become the dominant population of B-cells and plasma cells within the repertoire- akin to Darwin's 'survival of the fittest'.

The human intellect has appropriated these biological mechanisms to its advantage. The first instance of generating monoclonal antibodies entailed the immunization of mice with an immunogen that skewed their repertoires to contain plasma cells that possess high affinity for the antigen. These were then fused with special immortalized cell lines, that enabled the selection of fused cells, or hybridomas [13]. The next step in antibody discovery was akin to simulating natural selection within a test tube. Libraries of antibody fragments were fused with a surface protein of a filamentous phage and then

introduced to an antigen to gauge their binding affinity [14,15]. The phages were retrieved by their ability to infect bacteria and then reamplified. This cycle of selection continued by gradually decreasing the antigen concentration with each round to induce competition between phages, thus isolating only the most effective binding phages. 'Display' technologies all share a common feature of exploiting pools of organisms (or mRNA) that enable the genotype/phenotype coupling of antibodies (or antibody fragments), and driving binding competition between members of the pool using antigenic bait. Yeast or mammalian cells, for instance, compete for fluorescent antigen, whereas mRNA or phage compete for immobilized antigen. Moreover, natural antibody diversity mechanisms can be mimicked by employing strategies to slightly alter each expressed antibody between successive rounds, such as using error-prone polymerases to alter the DNA or shuffling light chains- variation and evolution by natural selection in a tube.

The natural selection of antibodies has not only occurred at the molecular level, but also at the organismal level, as evidenced by the structural differences between human, lagomorph, and avian antibodies [16,17]. One remarkable example is that of the camelid VHH antibodies, in which a germline mutation introduced a splice site that removes the CH1 portion of the antibody responsible for binding the light chain [18,19]. Despite this modification, the resulting heavy chain-only antibody has been preserved through the course evolution, possibly due to its ability to fit into smaller antigenic grooves owing to its smaller size whilst still retaining a respectable binding affinity. Certain shark species also produce single domain antibodies that appear to be functional and have been preserved through evolution [20]- "natural selection is daily and hourly scrutinizing, throughout the world, every variation, even the slightest; rejecting that which is bad, preserving and adding up all that is good".

It is remarkable that Darwin's theory, which he applied to the large animals of the Savannas and the small finches of isolated islands, so beautifully explains the evolution of tiny biological macromolecules during an immune response that form the basis of immune protection against microscopic invaders. Although we can only speculate what Darwin would have made of this, there is undeniable grandeur in the mechanisms vertebrates use to maximize the size of natural antibody repertoires and further evolve and amplify the best binding antibody. Despite the microbes' constant attempts to invade these beasts in accordance with the fixed laws of germ theory, from such humble beginnings, countless immune molecules, most efficacious and most protective have been and are being evolved.

## References

1. Darwin, C. (1860). *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life* (Vol. 2). London: John Murray.
2. Liu, Y. (2018). In Search of Darwin's Imaginary Gemmules. *Advances in Genetics*, 101, 87-114.
3. Darwin, C. (1868). *The Variation of Animals and Plants under Domestication* (Vol. 2). London: John Murray.
4. Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex* (Vol. 2). London: John Murray.
5. Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, 302, 575–581.
6. Safonova, Y., & Pevzner, P. A. (2020). V(DD)J recombination is an important and evolutionarily conserved mechanism for generating antibodies with unusually long CDR3s. *Genome Research*, 30(11), 1547-1558.
7. Wu, T. T., Johnson, G., & Kabat, E. A. (1993). Length distribution of CDRH3 in antibodies. *Proteins*, 16(1), 1-7.
8. Neuberger, M. S., & Milstein, C. (1995). Somatic hypermutation. *Current Opinion in Immunology*, 7, 248–254.
9. Rees, A. R. (2020). Understanding the human antibody repertoire. *MAbs*, 12(1), 1729683.

10. Gay, D., Saunders, T., Camper, S., Weigert, M. (1993). Receptor editing: an approach by autoreactive B cells to escape tolerance. *Journal of Experimental Medicine*, 177, 999–1008.
11. Tiegs, S. L., Russell, D. M., & Nemazee, D. (1993). Receptor editing in self-reactive bone marrow B cells. *Journal of Experimental Medicine*, 177, 1009–1020.
12. Goodnow, C. C., Crosbie, J., Adelstein, S., Lavoie, T. B., Smith-Gill, S. J., Brink, R. A., Pritchard-Briscoe, H., Wotherspoon, J. S., Loblay, R. H., Raphael, K., & Trent, R. J. (1988). Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*, 334(6184), 676-682.
13. Köhler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256, 495–497.
14. McCafferty, J., Griffiths, A. D., Winter, G., & Chiswell, D. J. (1990). Phage antibodies: Filamentous phage displaying antibody variable domains. *Nature*, 348(6301), 552-554.
15. Clackson, T., Hoogenboom, H. R., Griffiths, A. D., & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature*, 352(6336), 624-628.
16. Su, C., & Nei, M. (1999). Fifty-million-year-old polymorphism at an immunoglobulin variable region gene locus in the rabbit evolutionary lineage. *Proceedings of the National Academy of Sciences of the United States of America*, 96(17), 9710-9715.
17. Kaspers, B., & Göbel, T. W. F. (2016). The Avian Immune System. In *Encyclopedia of Immunology* (2nd ed., pp. 216-221). Academic Press.
18. Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., ... & Bendahman, N. (1993). Naturally occurring antibodies devoid of light chains. *Nature*, 363(6428), 446-448.
19. Nguyen, V. K., Hamers, R., Wyns, L., & Muyldermans, S. (1999). Loss of splice consensus signal is responsible for the removal of the entire CH1 domain of the functional camel IGG2A heavy-chain antibodies. *Molecular immunology*, 36(8), 515-524.
20. Greenberg, A. S., Avila, D., Hughes, M., Hughes, A., McKinney, E. C., & Flajnik, M. F. (1995). A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature*, 374(6518), 168-173.