JULIE WHITEFIELD AND MARTIN KENNEDY

Above the Threshold

JULIE WHITEFIELD

Above the Threshold is an art installation produced for the Art and Genetics Exhibition, 2017, held in the HD Skinner Annex, Otago Museum, Dunedin, New Zealand, 4-16 July, 2017. Held in collaboration with the University of Otago, the Dunedin School of Art and the Otago Museum, I would like to thank Dr Ruth Napper and Peter Stupples for their generous organisation and curation. This venture allowed me a nostalgic revisit to the early 1980s as a PhD student at Auckland University, where I first met Martin Kennedy in the Molecular Genetics Lab.

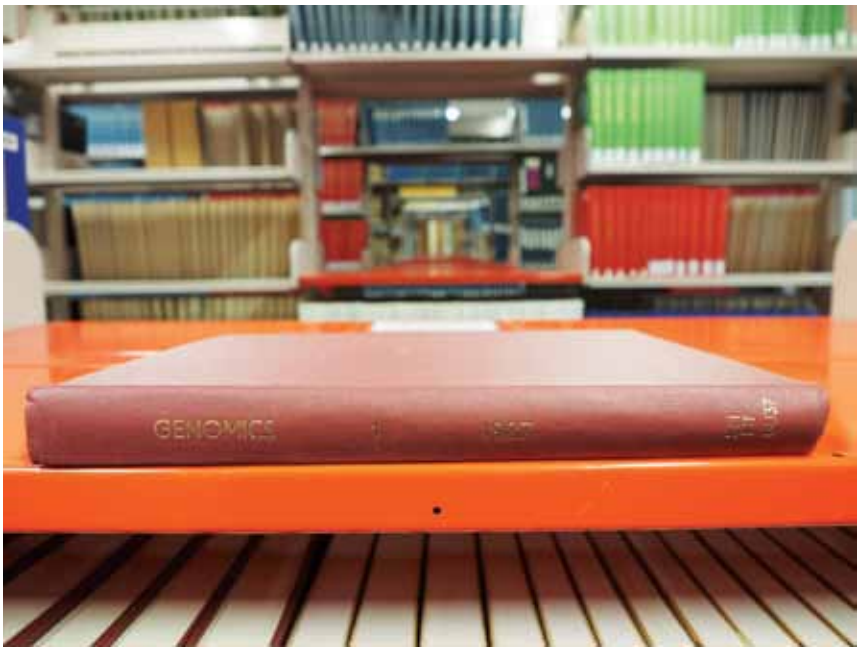


Figure 1. First publication of *Genomics* journal, 1987.

Whilst waiting to meet Martin, I ventured into the closest building at the academic institution to try and get the feel of the place back. The first thing I noticed in the University of Otago's Science Library was the lack of bound journals and racks of current journals on the main and first floors and the prevalence of computer screens and laptops. I had to climb to the second floor to feel within my comfort zone. Here I found some of the radical revolutions that have taken place in the field of genetics: the advent of Polymerase Chain Reaction (PCR), the Human Genome Project, Genome-Wide Association Studies (GWAS) and Next Generation Sequencing (NGS). I felt very out of touch.

Epigenetics, genomics, bioinformatics are all disciplines that have come about since the late 1980s. Martin's research is varied and incredibly complex. His group looks at how human genetic variation impacts on disease and its treatment. He is interested in patterns of genetic variation which underlie multifaceted human conditions, including psychiatric disorders, and how this variation impacts on the response to therapeutic drugs (pharmacogenomics) and particularly the occurrence of adverse drug reactions. Alongside these long-standing studies he has a couple of newer research topics. One is to explore intriguing structures called G-quadruplexes that form in DNA, and which appear to have a range of important biological functions. The other is to examine epigenetic changes triggered in the DNA by environmental factors.

The laboratory looked familiar, but processes seemed to be done on a much smaller scale and much of the work done by or on computers. PCRs are still performed, but most of the other techniques such as DNA extraction, primer production and DNA sequencing are outsourced. The days of hands-on laboratory work seemed to be nearly gone. It was nostalgic to see the sense of camaraderie in his laboratory alive and well in the form of a limerick competition, which it appeared that Martin won!! Tearoom discussions and the colloquial banter of staff and students is something I miss and enjoyed the chance, although temporary, to reclaim.

Martin introduced me to his world of bio-databases available on the internet. Incredible advances in DNA sequencing technology (NGS) and generous collaboration between scientists has allowed accessibility to volumes of sequence information and a variety of programs that allow analysis of this data. GTEx, gnomAD, GWAS, HapMap, genome browser, SNPs, all acronyms and concepts that were not technically possible during my research time in the field. I was also overwhelmed with the sense of the significance that statistical math plays in the role of these analyses. One of the mathematical tools Martin uses is the Manhattan plot.

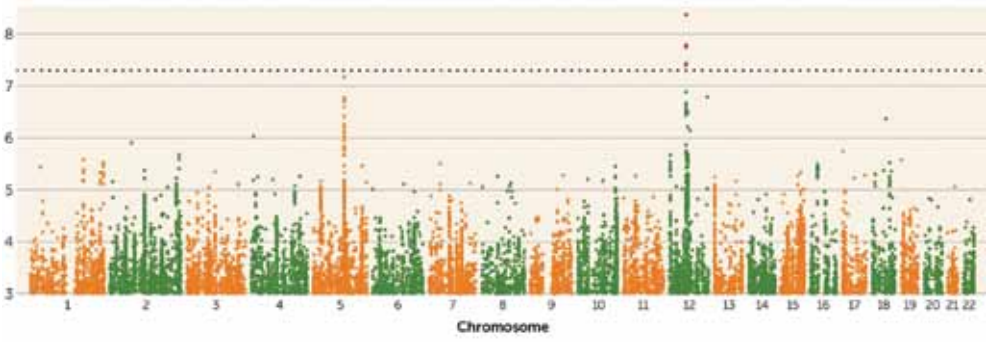


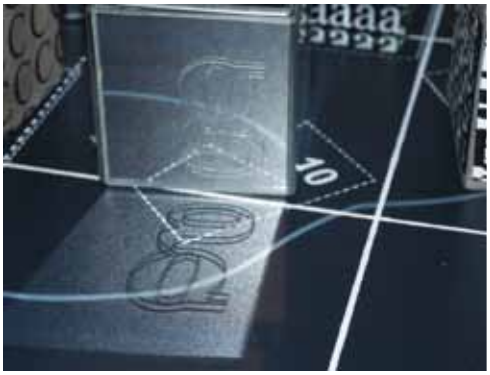
Figure 2. Example of Manhattan plot.[1]

A Manhattan plot is a type of scatter plot, used to display data where there are a large number of data-points. Each chromosome pair is displayed along the X-axis, with the negative logarithm of the association P-value for each single nucleotide polymorphism (SNP) displayed on the Y-axis, meaning that each dot on the Manhattan plot signifies a SNP. The strongest associations have the smallest P-values (e.g., 10−15), where their negative logarithms are the greatest (e.g., 15). Those points that lie above the threshold figure are deemed significant. The plot gains its name from its visual similarity to the Manhattan skyline: a profile of skyscrapers towering above the lower level "buildings" which vary around a lower height.

My first idea was to create something removed from the world of the computer, but still related to the overwhelming amount of data it produced and analysed. Another interplaying factor was the growing appreciation of the 3D architecture of the genome (epigenetics) playing a role in its functionality. However, the overriding purpose was to create something which was simple, accessible and referenced an everyday object to allow a connection out of the research laboratory. The data produced and analysed in the laboratory represents real people, mainly New Zealanders, mostly local to Christchurch. The resonance of place is lost in the data. The researcher is removed from the human element they are looking at. The aim was therefore to re-humanise the data.

The notion of creating children's building blocks as metaphors for SNPs emerged. Universally recognised as something to play with, but which also inform and hold knowledge, their tactility was

Figure 3. Close-up of the gridded "map" of Manhattan Island.



Figures 4a and 4b. Details of some blocks and placement on the grid.

an important feature in the aim to engage the viewer. Four of the faces were coded with letters symbolizing the four nucleotides present in DNA. The blocks relate to adult-sized hands and can be moved, are transposable, and able to be rotated in a column or row to produce a different sequence. Thus, they can mimic the SNPs that can occur in a DNA sequence at one position. The variability in architecture of DNA which can result from cytosine methylation and G-quadruplex formation can affect the relationship between the bases, between neighbours, between domains. They may not be related sequentially, but they are topographically. The 3D nature of the blocks also allows this reading.

The use of grids, graphs, charts and modular elements is a well-known idiom of scientific representation used to help understand and explain the inner workings of DNA. These abstract depictions are not so much appreciated for their objecthood, but more valued for their function of distilling and clarifying aspects of relevance into a symbolic form which can be analysed and interpreted. The simplified knowledge acts as a map to help navigate further from the known to the unknown.

> Maps are artefacts that help us make decisions, in so much as they visually organize data and information on a space; their aim is to make what they see comprehensible and usable, to bring it to our knowledge.[2]

In a quest for answers, maps, charts and plots highlight invisible links as seen through the eyes of the cartographer. In the case of the Manhattan plot, the cartographer sits between multi-disciplines: the statistician who created the mathematical boundaries of certainty and uncertainty, the medical practitioner who provided the diagnoses of a particular disease or drug reaction and the geneticist who sequenced the DNA and selected the DNA microarray platform to use. All are bound by rules with which to engage with the plot. The DNA variations highlighted are an objective and clinical portrayal of the genome, not a result of the emotional experiences of the subject. The evaluation of the information is the last link in the chain, the cognition of the observers' reading of the map.

This work serves to create a further link in the chain, distancing the viewer one more step from reality and allowing more open-ended interpretations with fewer rules. A gridded "map" of Manhattan Island depicts 23 squares bound by broken white lines. Each empty square space creates a tension of expectancy, a gap waiting to be filled. The blocks occupying the space and representing a chromosome pair, analogous to the X axis of the Manhattan plot, are modular but can be read through different planes. Each block is printed with the same information but in different patterns. The process of repetition in screen-printing and stencilling mimics the replication and production of contiguous overlapping sequences involved in sequencing DNA.

Despite using technology (power and electronic tools) to manufacture the blocks, and develop the stencils and screen prints, the assembly of the blocks and application of paint layers has been executed by hand. Unlike more flexible textiles, the hard surface of the medium density fibreboard (mdf) and the geometry of the block left little room for disguising errors: there are discrepancies in the squareness of the blocks and their exact size (due to over-exuberant sanding), the painting is not "clean" and the printing has been aligned by eye, not with the precision of machine manufacturing. This disturbance of repetition creates differences in the sameness of what is otherwise a very ordered and predictable progression. This deliberate act leaves a sense of the touch of the maker.

Figure 5. Installation view, detail, Art and Genetics Exhibition, 2017.

My image choices revolved around the letters symbolizing the four nucleotides present in DNA: adenine (A), cytosine (C), guanine (G) and thymine (T). In most cases where there is a repetition of a visual phrase, the sequence is ordered in a linear fashion around the block, with the colour blue referencing the blueprint of an architect's building plans and DNA as the blueprint of life. Some blocks contain part of the MEST sequence which includes the G-quadruplex structures worked on in Martin's group.[3] In other blocks, the complexity of mapping is reduced to banding patterns centred on a nostalgic inclusion of a Sanger-based sequencing gel image as source material. A microarray plate print was also included as a decorative grid. I did not complicate imagery with chemical structures or helical iconography. This choice was done intentionally to emphasize and expand the conversation of DNA code as a data source and allow a more open interpretation to the building blocks, allowing them to operate in an inter-disciplinary space of science and art. They could be associated with other genres of textual and schematically representative diagrams, the give-away being the repetition of A,C,G and T in an alphabet of 26 letters.

Binary pairing of Adenine and Thymine, Cytosine and Guanine were depicted on opposite faces on the block. However, flaws occurred in this plan based on my memory and time lapses in manufacturing. Rather than redo the blocks I decided to leave these blocks in as flaws in the scheme, alluding to mutation and the imperfection of data.

The upright shape of the towers of blocks reflects the figurative representation of a mapped chromosome and also references the DNA microarrays and DNA chips used in NGS where the DNA of interest or primers are annealed to a solid platform. This grounding allows the sequenced DNA to rise up from a gridded platform. The topological arrangement and subsequent recesses in the work allow parts of the sequence to be buried. The information is therefore both accessible and inaccessible, reflecting the folding of DNA to create neighbourhoods and domains of interaction which mimic the epigenetic factors involved in the architecture and subsequent reading of DNA. Indeed, the viewer can interact with the installation and change blocks around, creating a real-time element. This capacity serves as a visual reminder of the mutability and change that occurs over time.

By their nature and placement, the blocks also represent offices, apartments, workplaces, houses; homes of people. They are preloaded cultural symbols of habitation and signify humanity and the variance that is found within. The inherent rhythm of repetition in DNA and the layering of levels of information create a unique series of stories handed down from generation to generation. Each edition retains a part of the preceding version, combined with the new: a seamless narrative of existence without beginning or end. DNA reflects the identity of its custodian, yet each person is a "self", not only because of their biochemical markers of individuality but also because they are the repository of a unique set of experiences and thereby endowed with a unique set of choices. The

Figure 6. Installation view, detail, Art and Genetics Exhibition, 2017.

range of buildings and eras that are represented within Manhattan Island reflect the differences present in an individual's DNA: glass towers, multi-storey buildings, older tenement buildings. They also mirror the generous and open collaboration of scientific researchers from all over the world, which has allowed the data to be collected and be freely available.

The cohesion of place in geographical space provides a cohesion of groups and enables interaction between otherwise disparate elements. Linking the data back to a geographical element creates a sense of belonging, a resonance of place and identity. Although not in New Zealand and thus seen as elsewhere, it focuses the reading of the data to be more specific. Manhattan also represents a microcosm of the First World, a cultural, financial, historical and socially significant centre for contemporary life. How often do images of Manhattan buildings flash across the computer, TV and movie screen? Not only does Manhattan house the most densely populated borough of New York City, but the power and wealth contained in this area command some of the most expensive real estate in the world.[4] Aligning DNA blocks in this space creates a conversation between the value and power of DNA as a commodity[5] and that inherent in the likes of Wall Street and Fifth Avenue.

DNA does not operate on land, in a grid; it functions in its own liquid sea and has the ability to change structure in less than a human heartbeat. Martin encounters data of billions of nucleotides looking for several, even hundreds, which are altered in a genomic sequence and can be linked with

Figures 7a and 7b. "Sea of data" table, close-up and installation view showing interactive categorisation and sorting of letters by viewers.

a specific phenotype or disease. These large numbers become anaesthetizing. To try and visually capture the overwhelming and incomprehensible enormity of this data, a literal "sea of data" sits alongside the Manhattan plot. The four nucleotide bases are represented as three-dimensional letters heaped into a mound on a plain ungridded background. The background is reminiscent of a stainless-steel bench used for dissection, a sorting table. The laser cut letters are in different sizes, both upper and lower case, and in Times New Roman font, a font associated with the credibility and professionalism of scientific academic writing. Most of the letters are painted colours associated with the sea: blues, greens, blacks, and foamy whites. Red letters hide within the monotony of the nucleotides and represent the changed nucleotides being searched for; and although then inherently aligned with a disease and considered negative, they are also the bearers of hope for the future.

The mish-mash of letters has no order, no sequence and as such cannot be considered a coherent entity; it is waiting, in a temporal space, its potential unrealised. The amorphous and changing structure and repetition declare the sequences infinite, referencing the seemingly endless task of marrying particular SNPs to phenotypic traits. Nelkin (1996) describes DNA as "text without context, data without dimension."[6] This lack of order contrasts with the order inherent in the Manhattan plot, calling attention to variation within the sameness. It was encouraging to see the interaction of viewers with the disarray of letters. People innately tried to categorise and sort them.

Above the Threshold is derived from a schematic representation, the Manhattan plot. It sits alongside this notion in parallel. Reinterpretation of a scientifically graphical tradition creates a different narrative, creating a bridge to other viewers and other interpretations. Not bound by the objective regimes of 'total truth' and functionality, this work crosses boundaries of scientific and artistic practice.[7] Although art does not necessarily need to be informative, the placement of information in art creates 'in-between' spaces that modifies the information and allows it to sit on a threshold. This crossing-over of disciplines exemplifies the permeability of boundaries and barriers.

According to Brian Greene, a theoretical physicist, artists and scientists both strive "to figure out the deep truths of reality".[8] This human desire arises from the need to understand the unknown and reframe the results in an accessible form. Data is used like a lens to identify patterns which may be related to the search for the 'truth', to give meaning and context to behaviours and diseases we do not understand. We approach this 'truth' often as an asymptotic curve to the axis line, never quite intersecting. Martin works in this grey space, not the dichotomous duality of black and white. Identifying the many genetic risk factors is the first step in a daunting task of understanding how variability manifests in a particular phenotype, the aim being to reduce the size of the 'black box' between the two. His search for truth is still based around difference and sameness in a binary fashion, but the variables he encounters seem endless. He seeks precision. As an artist, I am allowed the luxury of embracing ambiguity. Moving away from a scientific domain allows an opening of possibilities, being able to 'weight' concepts and objects differently and not being bound by the laws of accuracy, rigour and reproducibility.

During this project, I learnt a lot about Manhattan and New York. I learnt a lot about the advances in technology in molecular genetics over the past 30 years and I learnt that despite all the analytical examinations and dissections of DNA on so many levels, DNA remains elusive, an extraordinarily elegant and eloquent molecule. The molecule itself and its implications remain much more than the sum of all of the perspective that the still frames captured.

MARTIN KENNEDY

Julie and I were friends and colleagues when we were both starting out on our scientific journeys as PhD students at University of Auckland in the early 1980s. We ended up having very different career paths, so when she approached me about collaborating on this venture with her, I was intrigued and excited.

For as long as the structure of DNA has been known (over 60 years), it has been used in art of many forms. I think it is the fascination with DNA as the key molecule which underpins and transmits life that makes it so alluring and prone to creative interpretation by artists. But those are precisely the reasons which attract me, as a scientist, to DNA. As an undergraduate student, I was hooked by a lecture on genetic engineering. I loved the mystique and possibilities that manipulation of DNA offered, particularly through processes such as molecular cloning and DNA sequencing. I have never lost that love of the molecule, which is a basic driver underlying my career, but even better is that the research my team and I do can impact on the understanding and treatment of disease, which makes the hard work very rewarding. Needless to say, the ability to rapidly sequence genomes with modern technology makes this research space more exciting than ever, and keeps fuelling my fascination with this molecule.

I love explaining how genome-wide association studies (GWAS) work and what they have meant for our understanding of many diseases and other human traits, including mental disorders and adverse drug reactions, that are a particular focus of our work. The Manhattan plot is an icon of early twenty-first century human genetics; the science, technology and massive logistics required to generate each plot is testimony to the power of globally collaborative research which is changing the face of medicine forever. I was delighted when Julie probed me about Manhattan plots, and got hooked on this as a key theme for her work. In the world of genomics, data visualisation and presentation is both an art form and a vital key to understanding, and I think she very quickly grasped this point. Indeed, it is clear that Julie's research and development on this project was aided by her prior training in genetics, and allowed her to understand and interpret new genetic knowledge in very sophisticated ways.

Using children's building blocks as metaphors for the DNA bases and variants (SNPs) was inspired, and seems almost obvious in hindsight – but so did the pivotal method of PCR when it was invented over 30 years ago, and everyone wondered why *they* didn't think of it after the inventor received his Nobel Prize! Building the Manhattan towers with these blocks and leaving empty spaces on the map are elegant ways of representing what we have come to know and what still remains to be found by GWAS. That the building blocks crafted are not perfect actually reflects biology – the very variation which contributes to human individuality arises from the imperfect nature of the processes which replicate DNA every time a cell divides. Her observations on the three-dimensional aspects of the towers and their components also allude to the importance of the three-dimensional architecture of the genome, and how our increasing knowledge of this very organised structure is helping to explain how many mutations and DNA variations contribute to disease. And last, but definitely not least, the red letters in the "sea of data" — although an artistic device in the work - is precisely how we often present mutations or variant DNA bases in printed representations of DNA sequence.

Julie observed that "the days of hands-on laboratory work seemed to be nearly gone". I would

comment that to look in my laboratory now is to see quite different work patterns to those of ten years ago, when all my students and staff would be engaged in "bench work", carrying out "wet lab" experiments rather than "in silico" computer-aided analyses on enormous data sets. But while generating and analysing big genomic data sets is our main *modus operandi* now, it is because we are in a busy discovery period as far as human genetic variation and its relevance to disease and health goes. The next big challenges will be mainly around understanding *precisely* how all those newly discovered genome variants contribute to disease, and to achieve that level of knowledge will require a return to the laboratory and many "wet lab" experiments using biochemistry and molecular biological methods. So, I believe these things are really only in abeyance, and the pendulum will soon swing back to large-scale biological experimentation in order to progress understanding in genetics.

When I was a child I enjoyed reading about early explorers and their adventures. It is pretty hard to find new parts of the planet to explore these days. However, I have always felt that scientific research is full of new frontiers, and it is a way in which we get to continue the role of those early explorers. Seeing Julie apply an artist's eye, and her early genetics training, to human genomics has been both inspirational and educational. Turning relatively abstract concepts into everyday objects such as building blocks and towers, and doing this with impressive craft and artistry, is a suitable tribute to, and novel interpretation of, the power of modern genomics to understand the human condition.

Dr Julie Whitefield BSc MSc (Hons), PhD (Auckland) 1987, completed a Bachelor of Fine Arts in 2003, graduating from Otago Polytechnic. As a multi-media artist she is involved in art-based community projects and group exhibitions.

Professor Martin Kennedy BSc (Hons) (Canterbury), PhD (Auckland) 1986, is head of the Department of Pathology, director of the Carney Centre for Pharmacogenomics and laboratory director of the Gene Structure and Function Laboratory, University of Otago in Christchurch.

1. Laramie Duncan, et al., "Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa." *American Journal of Psychiatry* (2017), https://doi.org/10.1176/appi.ajp.2017.16121402
2. Luigi Farrauto, "The Naked Notebooks," in *Mind, Maps and Infographics*, curated by Pietro Corraini (Italy: Moleskine SpA, 2016) 13.
3. Aaron Stevens, et al., "G-Quadruplex Structures and CpG Methyladtion cause Drop-Out of the Maternal Allele in Polymerase Chain Reaction Amplification of the Imprinted MEST Gene Promoter," *PLoS ONE*, Vol. 9 (2014), https://doi.org/10.1371/journal.pone.0113955.
4. Wikipedia, "Manhattan," *Wikipedia*, https://en.wikipedia.org/wiki/Manhattan (accessed 16 March 2017).
5. Yaniv Erlich and Dina Zielinski, "DNA Fountain Enables a Robust and Efficient Storage Architecture." *Science*, Vol. 355 (2017): 950-954. The use of DNA as a commodity to store digital information when computer space on physical planar objects has reached its limits is becoming a possibility in the near future. Erlich and Zielinski report perfect retrieval from a density of 215 petabytes (10^{15}) per gram of DNA. The cost and practicality of this process has still to be refined, but potentially affords huge storage potential in a minute space.
6. Dorothy Nelkin, "The Gene as a Cultural Icon: Visual Images of DNA." *Art Journal*, Vol. 55 (1996) 56-61.
7. Giorgia Lupi, "Data Humanism: The Revolutionary Future of Data Visualisation," *Print* (2017) http://www.printmag.com/information-design/data-humanism-future-of-data-visualization/ (accessed 12 February 2017). Giorgia Lupi describes the challenges of data visualization designers: "We have to find new languages, and explore how to convey knowledge and inspire feelings simultaneously with data. We have to discover how to be faithful to scientific accuracy while allowing space for exceptions to flourish. We have to bring data to life — human life."
8. Brian Greene, "Science and Art are Not Separate Disciplines," *Big Think* (2015). http://bigthink.com/videos/brian-greene-on-the-overlap-of-science-and-art (accessed 16 March 2017).