Genet. Res., Camb. (2006), **88**, *pp.* 191–196.
Printed in the United Kingdom

Book Reviews

doi:10.1017/S0016672307008531

Won for All: How the Drosophila genome was sequenced. M. Ashburner. Cold Spring Harbor Laboratory Press. 2006. 107 pages. ISBN 087969802 0. Price £11 (hardback).

"*There is no reason to switch.*"

This short book reminds us of the earliest critical decisions in the field of genome sequencing. It is well worth reading it together with two other books that tell other parts of the larger story of which the fly genome sequence is part. The first is John Sulston and Georgina Ferry's "The Common Thread" and the second is "The Genome War" by James Shreeve, a journalist who was given access to Celera. These three books are significant because the very public conflict between Celera Genomics and the academic sequencing projects to complete a draft sequence of the human genome is having long-term adverse effects on the perception of how academic biology works, especially in the US and more especially in the US government. How did the "Genome War" develop and might it have been prevented or made less damaging?

John Sulston believes that the main factor in not getting the human genome completed sooner as an academic project was the delay in starting the human genome sequence at the maximum rate that would have been possible in 1996. The MRC did not join with the Wellcome Foundation in approving financial support to Sulston for human genome sequencing at that point and the academic community did not wholeheartedly support the effort. While it seems unlikely that the human genome sequence could have been completed with slab gel sequencing technology anyhow, it might have been possible to have had more unanimous academic support for the way the genome sequence was being approached before a commercial competitor appeared. Does the available written history of genome sequencing explain what happened and how did it involve the fly genome?

This little memoir begins snappily with Craig Venter's announcement at a Cold Spring Harbor genome meeting in 1998 that he will lead a company to sequence the human genome and sequence *Drosophila* first to prove the method. The main weapon will be a new generation of much faster automated capillary sequencing machines developed by the Hunkapillers for Perkin Elmer. Gerry Rubin, leader of the Berkeley *Drosophila* Genome Sequencing Project, agrees to cooperate. Michael Ashburner was head of the European Bioinformatics Institute at Hinxton at the time and this book is mostly about the work involved in interpreting the whole genome sequence and the community of people who contributed. Michael helped organize an Annotation Jamboree at Celera where the first round of annotation on the fly sequence was done.

The interactions described here between the fly people and Celera are fraught. The *Drosophila* genome sequence has some commercial value and intense negotiation is required to ensure that it is made available to academic users without restrictions. However, on the day that Celera first make the *Drosophila* genome sequence available to NCBI it comes with constraints that had not been agreed. Michael, a leading advocate of free access to scientific information is the one who cries "Foul!" and recruits others to force Celera to remove the restrictions. Michael refers to Shreeve's book for proof that the company lawyers had indeed been testing the water with this ploy. It was important therefore that Michael ensured the rebuff.

The title "Won for all." refers to the spirit of challenged brotherhood that developed between the main academic genome sequencing centers in response to Celera. Michael is toeing the standard academic line here but his attitude to the public genome sequencing projects is of equal interest and this is more ambivalent than his book title suggests. What we tend to forget about The Three Musketeers is that d'Artagnon was not in fact one of them but a brasher person with the same objectives. The person who emerges as d'Artagnon here and in the Shreeve book is not so much Craig Venter as Gene Myers, the computer scientist who led the assembly of the genomes for Celera. He was helped by an emeritus Hamilton Smith who made the clone libraries.

Myers had several years earlier co-authored a suggestion that the human genome could be assembled without first building a complete physical map of overlapping clones. Instead, carefully sized random genome fragments sequenced at opposite ends to generate "mate pair" sequences at a known distance apart would provide enough positional information to order the full genome sequence assembly. Michael quotes Phil Green, a close colleague of Bob Waterston and John Sulston who were the main genome sequencers at that time, in his authoritative and not entirely gentle rejection of this suggestion for a change of approach; "There is no reason to switch.". The fly genome assembly showed that Myers had been substantially correct; this would be the main way to sequence genomes in the future. The Celera genome assembly was validated by comparison to the 20% of the fly genome that had already been sequenced mainly by Rubin's project at Berkeley. Remaining gaps were closed and the genome finished at Berkeley. Myers who is now back in academia, in Gerry Rubin's new research campus at Janelia Farm.

While much of the argument of "Won for all." is directed against commercial motivations and in favour of open scientific information, criticising Celera seems a little like flogging a dead horse at this point in time. This book would be better if Michael had told the full story of the slow steps to the sequencing of the fly genome and its wider relevance from his own perspective even if others might disagree with it. Instead he avoids discussing the bigger issue of the human genome sequence and we have to take hints from his somewhat cryptic pointing out of inflexibility in the public sequencing projects. Why is Michael still criticising only Celera? Would a more even-handed presentation have forced him to embarrass his friends by discussing other instances where the academic sequencers got it wrong?

A more complete version of this book would begin in 1989 when the clone map of the worm genome was completed. Jim Watson suggested at his point that the fly genome rather than the worm genome should be sequenced. Watson tried to get support for this view within the British MRC and, as interviews carried out by Georgina Ferry show, Aaron Klug, then head of the LMB, was sufficiently convinced to attempt to argue the matter with John Sulston. Watson understood that genome sequencing could not capture the imagination and the support of biologists and go ahead at the maximum pace unless the fly genome rather than the worm genome was done first. It was already clear in 1989 that the fly sequence had much greater similarity to human sequence than the worm sequence did. Also the greater similarity between fly and vertebrate in the control of organ formation and in *Hox* gene conservation and function, for example, would make the fly a much more useful choice than worm to advance the argument for comparative genomics and thereby human genome sequencing. What happened instead was that, by 1996, the really exciting sequences had been produced by individual labs working on fly and mouse and human. The worm genome sequence, the only product of large scale animal genome sequencing was pretty dull and we were not excited about spending vast sums on the human genome.

Homologs of many important human genes found in the fly had not been detected in the worm, because the greater sequence divergence made it impossible to see conservations that were not strongly expected. On the other hand the really spectacular examples of sequence and functional conservation between fly and human genes were the ones that could have altered the fortunes of genome sequencing if the fly had been sequenced first. In 1995 Walter Gehring's laboratory in Basel showed that the fly *Eyeless* gene or its human homolog *PAX6* would produce ectopic eyes on legs and other parts of the body in flies. The fly ectopic eyes provided compelling evidence to human and vertebrate geneticists of the potential of comparative genomics. If the *PAX6* homolog and other fly homologs of human genes had been identified in a fly genome sequence being done in Cambridge at that time then genome sequencing would have shared in the credit for this and many other discoveries. Celera rubbed salt in the Sanger Centre's wounds by correcting the mistake, sequencing *Drosophila* as a public service. The fly genome sequence made it clear, above all, just how much better the reputation of genome sequencing could have been at the crucial moment in 1996 if the fly had been sequenced first.

Michael in this book recognizes the similarities between the human genome sequence and the Apollo project. If the Sanger Centre had sequenced *Drosophila* in the early nineties the effect of all those exciting and dramatic conserved gene homologs arriving together would have been electric. Jim Watson now likes to ask, "What has happened to British science?". The Sanger Centre was the leading sequencing center at the time and they lost a great opportunity. Reading "The Common Thread" suggests that part of the problem may be the narrowness of traditional British PhD training that leaves people at later stages in their careers struggling to make balanced decisions over a wider range of biology.

It is worth reconsidering the worm-fly genome sequence debate, brief though it was, because the genome sequencing Apollo project is not yet finished. We now probably have to make billion dollar investments in projects to advance genome sequencing to the point where it will be affordable for each of us to have our genome sequenced as a routine part of clinical practice. Once again we must choose the most worthwhile biological target for a huge project and

like the worm-fly choice the decision will affect the progress of academic biology. One of the most disturbing aspects of the history of the worm-fly genome decision is John Sulston's claim that they went ahead with worm partly because the clones were already available. It was the challenge of scaling up, of developing the large scale sequencing methodologies on easily available DNA that excited them more than the choice of biological target.

LIAM KEEGAN

Hyde Park, Killucan, Co. Westmeath, Ireland

MRC Human Genetics Unit, Western General Hospital, Edinburgh, EH4 2EH

doi:10.1017/S0016672307008543

Population genetics, quantitative genetics and animal improvement: papers in honour of William (Bill) Hill. *Philosophical Transactions of the Royal Society B* (2005) **360**. Edited by B. Charlesworth, P. Keightley & P. Visscher

This volume was published as a (surprise, I believe) tribute to Bill Hill[1] on the occasion of his 65th birthday. As anyone in quantitative or population genetics should know, Bill has made numerous important contributions to the subjects in the title of this volume over the last 40 years. Bill must now be considered the "eminence grise" of quantitative genetics – a term he is bound to object to due to its connotations of established and unquestioned expert (and age!). Bill would never like to be unquestioned and still relishes a good debate; nevertheless, this collection reflects the breadth and depth of his contributions to these areas. Written by experts in their own right, the contributors include former students and current collaborators. Many of the contributors have worked or passed through the Department in Edinburgh where Bill has spent the majority of his time and a few have not.

The collection displays the vigour in quantitative genetics and related areas today. As the editors point out in their introduction, quantitative traits, influenced by a number of genes and environmental factors, include most of the important traits of evolutionary, economic and medical importance. For a while the molecular genetic revolution put quantitative genetics in the shade. However, in recent years it has become obvious that it is only possible to dissect complex traits, to understand how gene action and other factors control trait variation, to utilise quantitative traits in artificial breeding or disease prediction and to predict the outcome of selection or other interventions on such traits by combining the tools of genomics and molecular genetics with sophisticated models of quantitative genetics. The papers gathered here represent just some of the most important areas of development in genetics and all relate in some way to contributions of Bill Hill.

I assume the contributors were given relatively loose guidelines on the nature of their contributions as the structure varies from paper to paper. However, the papers fall broadly into two categories: authoritative overviews of a particular subject or more focussed novel contributions. All the papers are worth reading and it is unfair to pick out only a few for mention – however, I'll do it to give a flavour of the content of this volume. The collection opens with an excellent review by Toro and Caballero (2005) of genetic diversity in subdivided populations. The major focus and relevance of this contribution is to species under domestication or conservation and the authors review estimation of molecular and phenotypic diversity, the balance of within and between population diversity in the context of conservation and the genetic management of subdivided populations. This paper should really be obligatory reading for all those involved in conservation – it is both scholarly and comprehensive (within the limits of the space allowed it) and sets the tone for the remainder of the volume. In a similar vein of the scholarly overview are the contributions of Wang (2005) on the use of marker data to estimate effective population sizes and that and of Thomas (2005) who reviews use of marker data to infer relationships between individuals.

The nature and maintenance of quantitative variance is another fascinating subject that attracts the attention of experts in this volume. Johnson and Barton (2005) review the apparent incompatibility between the existence of substantial genetic variation for most quantitative traits and the fact that such traits can often be shown to be under stabilising selection that under simple models should act to purge the population of variation. They look at attempts, as yet not fully successful, to reconcile the facts, again providing a scholarly and readable review. Mackay and Lyman (2005) review understanding of the genetic control of *Drosophila* bristle numbers, archetypal quantitative traits, once thought of as simple traits under additive genetic control and subject to stabilising selection. Not only does this review reaffirm the value of studying this apparently simple model, but it also demonstrates that these traits are anything but simple. Amongst other things, the work summarised demonstrates the importance of pleiotropy (also a conclusion reached by Johnson and Barton), and the major roles played by complex interactions: gene by gene (i.e. epistasis), gene by sex and gene by environment. Thus theoretical and other studies of the maintenance of genetic variation need

[1] I have never heard anyone refer to him as William except in jest or at one of those daft formal occasions. Perhaps he has a maiden aunt who still refers to him as William?