



Conclusion

The heavy scholarly and media focus on the determination of the human reference sequence, popularly portrayed as the sequencing of ‘the human genome’, has had the effect of limiting the public perception of what constitutes genomics and its history. This perception concerns the characteristic practices, products and organisational configurations of genomics, and also locates genomics in a distinct era that closed with the ‘completion’ of the human reference genome in 2003. Genomics has become synonymous with a narrow set of practices and events associated with the creation of a reference sequence, and chiefly that of one species, the human. The industrial forms of production that this reference sequence required, as well as the possibilities it opened in biomedical research, has established rigid boundaries between what is conceived as pre-genomics, genomics and post-genomics. This periodisation foregrounds discontinuities and complicates any possible connection to be made between pre-genomic history and post-genomics. Throughout the preceding seven chapters, we have challenged this limited canonical view and argued that beyond *Homo sapiens* and the practice of large-scale sequencing, the historical vistas of genomics expand considerably.

In this final chapter, we reflect on the broader historiographical implications of our challenge. One of the consequences of the canonical view of genomics and its narrow historical lens has been that academic and policy appraisals of the nature and role of reference genomes treat them as

isolated objects preserved in aspic. Instead, we have uncovered the dynamics of genomics as well as the changes arising from—and happening to—reference genomes over the course of: their establishment as scientific objects; the efforts to produce them; activities to refine, improve and enrich them; and their connection to related resources built using them. Our presentation of different communities of genomicists with distinct, historically-specific mechanisms of inclusion and exclusion and the active role that these communities played in configuring the affordances and ontological status of each reference genome, has been crucial for our dynamic approach to genomics and its history.

The reflections in this final chapter help us to elaborate on the main conceptual payoffs of our analysis, namely the portrayal of the reference genome as a dynamic and generative entity that shapes our understanding of the past, present and future prospects of genomics research. We outline this using our key distinction between post-genomics and post-reference genomics and assess what this differentiation can offer us analytically in considering the question of research translation. Finally, we close with a discussion concerning how our multi-species approach and emphasis on communities of genomicists as historical actors affects the historiography of genomics. With this, we attempt to marry specificity with being able to make more global claims about genomics and its history.

8.1 THE NEVER-ENDING FRONTIER: QUERYING THE LIMITS OF GENOMICS AND CHARACTERISING PROGRESS WITHIN IT

An ideal of completeness and comprehensiveness has guided some of the leading promoters of genomics. What constitutes completeness and comprehensiveness has been, though, a continually receding horizon. Understanding the nature of the reference genome and the multidimensionality of webs of reference following the production of a reference sequence, as detailed in Chap. 7, helps us to appreciate why. Even if it is possible to determine end-to-end sequences without gaps—as it is now for humans (Nurk et al., 2022)—there can be no absolute and final way of apprehending and characterising the variation in and of a particular type, such as a species. The goals shift as the available data and knowledge grows and new research aims are developed. A surplus of potential representations and instantiations of variation becomes available to genomicists and

other researchers: the *variational surplus*. Variation is the measured or measurable differences of particular parameters within a defined type of object or process. In genomics, variants are detected by comparison of novel data with reference sequences or other standard resources. This variational surplus provides a plethora of potential routes through—and maps of—the variation, by which researchers can pursue their aims. There can therefore never be convergence on a final standard or ultimate set of linked standards.

One of the most compelling ways of interpreting the shifting frontier of what constitutes completeness or comprehensiveness is to consider that genomics manifests a particular open-ended version of what Hasok Chang has articulated as “epistemic iteration”. This concept, and its particular features, was drawn by Chang from his studies of the establishment of standards for the measurement of temperature and development of thermometers across eighteenth and nineteenth-century physics. Chang defines epistemic iteration as “a process in which successive stages of knowledge, each building on the preceding ones, are created in order to enhance the achievement of certain epistemic goals” (Chang, 2004, p. 45). He emphasises that his characterisation of the key features of epistemic iteration, while abstract, cannot necessarily be conceived as general or universal, even across the physical sciences, let alone the biological. We do not intend to risk plunging into the deep waters—around which Chang has posted warning signs—by merely transposing or applying epistemic iteration as developed in the context of thermometry, to the establishment and development of reference resources in genomics. Instead, by assessing the historical development of genomics, we adapt this conceptual framework. In so doing, we intend to shed light on genomics and also examine how to extend epistemic iteration to domains that appear quite different to the precision measurement of physical parameters.

There are a number of features of epistemic iteration that Chang identifies. A “correct answer” may not be knowable. Different stages need not feature the same knowledge production processes, nor be reducible to prior stages (Chang, 2004, pp. 45–46). What guides the process of iteration, according to Chang, is an “imperative of progress” judged against certain epistemic virtues and values (Chang, 2004, p. 44). Furthermore, although there is evident conservatism based on a “principle of respect” for prior standards (Chang, 2004, p. 43), this manifests in a “pluralistic traditionalism” in which “each line of inquiry needs to take place within a tradition, but the researcher is ultimately not confined to

the choice of one tradition, and each tradition can give rise to many competing lines of development” (Chang, 2004, p. 232).

In his discussion of thermometry, Chang details debates on the establishment of fixed points around which to base the temperature scale, the choice of substance (e.g. mercury or air) to incorporate into thermometers, the establishment of a theory-based absolute temperature and attempts to operationalise this by connecting it to concrete measurement methods. He demonstrates that some form of grounding on assumptions or imperfect empirical observations is necessary. Crucially, the improvement of standards—as evaluated against epistemic virtues, values and goals—often occurs through self-correction and enrichment, by building on and superseding prior standards.

There are some basic analogies between Chang’s discussion of the development of thermometry and the history of genomics and reference genomes. The reference genome is indeed, at any one time, a fixed point, a contingent result of consensus. However, over time it changes; no reference genome, at least yet, has attained the near-permanency of the Celsius and Fahrenheit scales. The choice of thermometry substance is analogous to the selection of the source material to be sequenced in a project to determine a reference genome. In genomics, though, rather than measurements and arguments being conducted by a community around a material, the material itself is a community product: a result of opportunity and availability (pig genomics), the prior history of the genomicists involved in the sequencing effort (the use of the S288C strain of yeast) or an attempt to represent, quasi-metaphysically, the species in question (human genomics).

What, though, restricts the stipulative freedom when producing and presenting a reference genome? What is there to stop it from being arbitrary? In line with the conservatism of the processes of producing temperature standards, reference genomes must be consistent with previously-established antecedents and exhibit improvement according to metrics of validation and evaluation that allow comparison of quality. Robust processes for ironing out sources of error (e.g. through deep sequence coverage, as well as statistical and computational means) are especially important when post-hoc detection of ‘errors’ may not be possible, and where the status of something as an error may itself be questioned. Epistemic goals are crucial in shaping this iterative process; indeed, we can identify what a particular community is seeking through the metrics it uses to validate new versions of reference resources. For instance, an

abstract idea of completeness and universality underlay the production of the human genome, whereas more specific agricultural and immunogenetic motives were behind the determination of the pig genome. Here, it may be observed that for partial or whole-genome sequence assemblies, using the quality of the assembly as some kind of context-free criterion, without reference to specific applications, may inhibit the use of it for other, translational purposes, and therefore complicate the development and usability of reference resources across communities.

It is important to note that epistemic goals in genomics are not merely subordinated to widely-held standards of quality or completeness. Throughout the history of genomics, different epistemic goals have motivated genomic research and data generation beyond just the creation of gold-standard reference genomes. And even for the creation of reference genomes, we have shown that maximising their completeness and quality according to certain metrics has not always been the sole or overriding concern of those promoting and conducting genomic projects. We have observed something distinctive about post-reference genomics, though, in that epistemic goals tend to shift towards the development and exploitation of reference resources built on and linked to the available reference genomes. These post-reference genomic resources characterise different forms of variation within the overall potential array of variation that can be apprehended and captured for a given species or across different species. Such aims to capture variation in this way existed before the advent of reference genomes, but once reference genomes are created, they present possibilities and opportunities to do this kind of work, ones that may not have been practical or conceivable before.

In the open-ended epistemic iteration characterising genomics, we therefore see the exploration of a particular *variational space*, by way of the creation of new genomic resources (data, materials, tools and infrastructures) that are based, in some respect, on the reference genome. This epistemic iteration constitutes a radiation from the fixed point of the reference genome, rather than a convergence to a fixed point as in Chang's thermometry. This explorative radiation is often conducted by a wider array of actors than were involved in the creation of reference genomes. It is shaped, though, by the initial conditions that are set by the processes by which the reference genome is produced, and is subsequently developed. In other words, the room for manoeuvre in post-reference genomics is shaped by the historicity of genomics: by the affordances and

representativeness that different communities of genomicists envisioned and enacted.¹ This is why the inclusion or exclusion of particular communities in the production of a reference genome is so important.

8.2 A DYNAMIC VIEW OF REFERENCE GENOMES AND THEIR ROLE

Throughout the book, we have shown how the processes and differential involvement of particular communities in the generation of reference genomes affect their nature and exploitation. Typically, criticisms of reference genomes within the life sciences and philosophy of science focus on matters related to the extent to which they represent or stand-in for their target species in meaningful ways (e.g., Ballouz et al., 2019; Barnes & Dupré, 2008; Rosenfeld et al., 2012; Tauber & Sarkar, 1992). We have argued, however, that the question of what the reference genome represents, and the identification of alleged deficiencies in the processes of abstraction, misleads us by directing attention only to the reference genome as an object or end in itself. When its role as an active foundation for the seeding of webs of reference is considered, the ways in which reference genomes are produced becomes pertinent to appreciating their infrastructural role, and not merely their representative one. These ways, we have shown, include the thicker array of practices and configurations involved in the production of a reference genome, and not just the determination of a string of nucleotides and the absences and presences in these.

The webs of linked reference resources built on and around a reference genome, in turn, feed into the ongoing development and context of use of the reference genome to further seed explorations of variational space. The reference genome is therefore a dynamic entity, shaped and reformed by the very processes of production that generated it, and by the webs of linked resources it has helped to create. Later in this section, we reflect on the ontological implications of these dynamics for reference genomes.

¹This is congruent with Soraya de Chadarevian’s observation of the dynamic relationship between the setting of standards and exploration of human variation in the histories of cytogenetics and genomics. She has noted that “[t]he search for variation, then, seems to be built deeply into the study of heredity. Yet how variation is interpreted—as variation on a theme or deviation from a norm, in a hierarchical or inclusive manner—and how it is acted upon, is a matter of interpretation and historical contingency” (de Chadarevian, 2020a; see also 2020b).

First, we pursue some suggestions about the type of object that reference genomes constitute, or have been thought to constitute.

Leading figures in human genomics were adamant both before and after the publication of the human reference sequence that the resulting object would not constitute a “normal genome” in any respect. In 1989, for example, Victor McKusick, the co-founder of the journal *Genomics* (Chap. 3), emphasised that it was “well recognized by geneticists, that there is no single normal, ideal, or perfect genome”. Interestingly, this was stated in justification of the idea that “the DNA can come from different persons chosen for study of particular parts of the genome. Such an approach is consistent with that of most biologic research, which depends on a few, and even on single individuals, to represent the whole”. After all, if the reference genome was not presumed to be normative in some way or another, then why should it matter what it represented? McKusick did not, however, suggest a completely arbitrary basis for the reference genome. Writing more than a decade before its accomplishment, he argued that the DNA would need to come from actual human beings, and its assembly would be guided by prior standards such as maps, with the reference sequence constituting “the ultimate map”, and validated according to other procedures to assess its quality and coverage (McKusick, 1989, p. 913).

Lisa Gannett (2003, pp. 179 and 182) identifies a range of positions on the idea of a “normal genome”, from David Hull and Elliott Sober’s “outright rejections of the notion of a normal genome and any treatment of genetic variation as deviation” to “the idea of a single genetic norm for the species from which all variation is deviation”. Advocates of the latter position appeal to evolution or adaptation to the environment as the basis for such a norm. Within this variety of views, McKusick’s contribution intended to present the reference genome as a kind of standard that abstracted from the genomic variation of the species but was not supposed to represent either the most common or the ‘best’ genome. To the extent that it is accepted as ‘normal’ by a community of practitioners—from genomicists involved in its production to other life scientists—it is a stipulated standard.

This conception of the human reference genome was to change. Writing a brief reflection on the tenth anniversary of the February 2001 draft sequence publication, with the benefit of the resulting knowledge gained about the human genome, Maynard Olson offered the view that “[a] model for human genetic individuality is emerging in which there actually

is a ‘wild-type’ human genome—one in which most genes exist in an evolutionarily optimized form”. He argued against this normative view on the grounds that “[t]here just are no ‘wild-type’ humans: we each fall short of this Platonic ideal in our own distinctive ways” (Olson, 2011, p. 872).

In his interpretation of the human reference genome, Olson—who played a crucial role in the mapping of the yeast *Saccharomyces cerevisiae* and devised tools to map larger genomes (Chap. 2)—referred to a particular concept, the wild type. This concept has been and is still used in medical genetics concerning a gene or functionally-relevant sequence that is not associated with a manifestation of disease or disorder. It therefore presumes that there are functional and non-pathogenic forms of genes. The wild type here is defined negatively, as not possessing certain forms of variation that would render the sequence non-functional or pathogenic. Since this is a function-first definition, what constitutes a wild type is not evident from the sequence itself: whatever deviates from the functional criteria used to assess the presence or absence of the wild type form of a gene or sequence is deemed to be a “mutation”.

There are other meanings of wild type that have been used in the life sciences, dating back to the early nineteenth century. At the outset of the twentieth century, a variety of interpretations of wild type flowered. It became applied by William Bateson, for instance, to organisms that exhibited a “normal body” as a result of experiencing “normal development”, as judged against the evolutionary history of the species. The wild type was therefore healthy and well-functioning, and a baseline against which variants could be assessed as beneficial or harmful. This was very much in line with the normative medical genetic version of it (Holmes, 2017; on normal development, see Lowe, 2016).

With the advent of what became known as ‘classical genetics’ in the laboratory of Thomas Hunt Morgan from the second decade of the twentieth century onwards, the wild type came to designate not just strains, individuals and genomes that represented the ‘normal’ as seen in nature, but also particular genes without evident mutant characters. So, for *Drosophila*, the wild type could refer to organisms with two symmetrical wings, the standard red eye colour, or other characteristics.

In this approach—that became prevalent in genetic experimentation—an organism or strain may be deemed to be a wild type, provided the characteristics pertinent to what is being investigated were themselves wild type. In this way, these characteristics serve as a baseline against which

deviations from the wild type—variation—can be apprehended and then interpreted. This shift enabled the articulation of genes as difference makers whose effect could be discerned not by the presence or absence of particular characters or traits but through comparing observable variation to a standard. By this point, wild types could not be considered to be wild, though they were supposed to stand-in for nature in the laboratory, and thus function as a correlate within the laboratory of the nature outside.

This assumption that laboratory wild type strains were supposed to constitute a particular reflection of standard traits and provide a means to apprehend and measure variation outside the laboratory came under devastating attack by neo-Darwinian ‘Modern Synthesis’ theorists in the mid-twentieth century. This critique highlighted the limitations of some of the programmes of research conducted using wild types, and undermined their conceptual basis. The wild type endured in the life sciences, however, as embodied in “standard lab strains of experimental organisms [...] [that] operate as controls to measure variation in model organism systems” (Holmes, 2017, p. 15). Indeed, the criticisms of the use of laboratory wild type strains also echo many of those levelled at the use of a small number of highly-standardised model organisms across biological research (Table 8.1).

Neither wild types nor model organisms account for the extent of natural variation. The very qualities that make a model organism useful for laboratory-based research also make them quite unlike even their wild cousins of the same species. Furthermore, the extent to which they possess the representational scope to capture biological processes and phenomena that occur in different species has been questioned (e.g. Bolker, 2017).

Table 8.1 The main model organisms used in biological research. Table elaborated by James Lowe

<i>Escherichia coli</i> (Bacteria)
<i>Saccharomyces cerevisiae</i> (Brewer’s and baker’s yeast)
<i>Arabidopsis thaliana</i> (Thale cress)
<i>Caenorhabditis elegans</i> (Nematode worm)
<i>Drosophila melanogaster</i> (Fruit fly)
<i>Danio rerio</i> (Zebrafish)
<i>Mus musculus</i> (Mouse)
<i>Gallus gallus</i> (Chicken)
<i>Xenopus laevis</i> (Frog)

Much of the recent concern over the translational gap between laboratory research and the clinic—e.g., relating to the development of new pharmaceutical products—has focused on the panoply of differences between laboratory workhorses such as the mouse *Mus musculus* and the humans who are supposed to benefit from such research (e.g. Garner, 2014).

Philosophical responses to such criticisms of the nature and use of model organisms have focused on their role as intensive hubs of resources concerning all aspects of the biology of the model organism species, which therefore function as a well-characterised basis for the generation of comparisons and the apprehension of variation across species (e.g. Ankeny & Leonelli, 2011; Leonelli & Ankeny, 2013; Ankeny, 2007, pp. 49–51; Leonelli, 2016, pp. 18–24, 145–148). Drawing on Rachel Ankeny’s analysis of work on *Caenorhabditis elegans*, Lisa Gannett has observed that, like model organisms, reference genomes constitute a kind of descriptive model, in that they instantiate an abstraction that is used as a foundation for explanatory questions (Ankeny, 2000; Gannett, 2019). In this sense, they should be assessed in terms of how they ground further research—as infrastructures—rather than on the extent to which they alone sufficiently represent the genomic variation of a species or sub-species.

One criticism within genomics itself concerning the utility and representativeness of reference genomes is that they act as type specimens: reference samples that taxonomists use “to define the general class by example, often for a species”. Reference genomes and type specimens share an “idiosyncratic” nature, in the sense that “[t]he data and assembly that made up the reference sequence reflect a highly specific process operating on highly specific samples”. This means that, even if a reference genome is a useful and “good” type specimen of its target species—which some critics admit for the human reference sequence—it cannot adequately reflect the variational landscape of that species in nature (Ballouz et al., 2019, quotes from pp. 1–3).

How apt a designation is this for reference genomes, and what would interpreting them as type specimens mean for understanding the nature and function of reference sequences and other genomic reference resources?

Type specimens are defined and used in the fields of taxonomy and systematics as standards around which practices of classification, and apprehension and cataloguing of variation, can operate. In taxonomy and systematics, type specimens are material instantiations of an organism, on

which the classification and name of a given type—such as a species—is anchored. This is vital for the enterprise of cataloguing and identification, and detailed specifications of different versions of type specimens have been developed by different communities. These kinds of designations, as well as the practices and rules governing them, have changed over time and also vary according to the kind of organism concerned, for example between animals and plants.

The use of type specimens in taxonomy has not been uncontroversial. It is intriguing how the questioning of their role has reflected some of the criticisms of reference genomes. For example, George Gaylord Simpson's critique of type specimens in the 1930s echoed the concern with how well they captured relevant variation to represent the type (Witteveen, 2018). Type specimens and reference genomes are indeed comparable, as both are fixed points of reference, at any one point in time. The representativeness of them in terms of biological variation is circumscribed, but they both enable variation to be apprehended, articulated, measured and recorded.

However, we emphasise that they are fixed points of reference *only at one point in time*. As the philosopher of biology Joeri Witteveen has noted (Witteveen, 2016), type specimens are not absolutely fixed as primary referents to particular species. They are, though, far less changeable than reference genomes have proven to be. We may speculate why this is the case. Possible reasons include the fact that reference genomes rely on already-designated species, and that they have a wider range and ever-changing set of epistemic goals that motivate continual iteration towards them. Furthermore, they have always been in digital form, allowing different versions to be designated and referred to far more easily. Reference genomes may offer a fixed point of reference, but serially rather than perpetually. By engaging with their historicity and the motivations of the communities of genomicists that created them—as we have done throughout the book—we can capture changes in their nature as references and as standards.

To introduce our assessment of that, we return to the determination of the human reference genome by the public and charitably-funded International Human Genome Sequencing Consortium (IHGSC) and the production of a whole-genome human sequence by the company Celera Genomics. They had different ways of generating their genomic data. Crucially, they also had different aims for the eventual product, which conditioned the strategies they pursued, but also their conception of the

objects they were creating. The IHGSC aimed to release, into the commons, a record of the ‘Book of Life’, *the* genetic code of the human species. This universalist view of the human genome was buttressed with data that indicated that DNA sequence similarity between humans was 99.9% and therefore far closer than in other species. Therefore, it did not seem to matter that the selection of donors was largely arbitrary, conducted through a newspaper advertisement (Chap. 4). IHGSC members argued that it was unnecessary and meaningless to use DNA from people of different ethnicities and sexes, as the differences in the DNA of humans across the globe were minimal.

Celera’s business model, on the other hand, was based on the identification and analysis of sequence variation. They wanted to sell that data to companies who would find it useful, for example in the development of diagnostic tools or therapeutic drugs. Later, they would try to exploit that data themselves for these purposes (García-Sancho, Leng, et al., 2022). Their emphasis was therefore on difference, rather than commonality or universality. Both efforts produced a comprehensive representation of the human genome, albeit one was a publicly released ‘official’ reference sequence, and the other was only available in full behind a paywall. Historians have already observed that these can indeed be regarded as two separate objects, because of the differential processes and configurations that went into producing them: Celera’s whole-genome shotgun approach and the IHGSC’s choice to construct physical maps and use these to help put the sequence together (Chap. 4; Bostanci, 2006).

Beyond that, we note that they constituted different forms of representation. For the IHGSC, their reference genome was representative of the species in the sense of faithfully depicting the genomes of humans across the world, except for a few minor and insignificant differences. For Celera, their genome was able to *stand-in* for the human species without substantially representing or reflecting its totality or diversity. At an event in August 2001, Gene Myers, a leading bioinformatician who worked at Celera from 1998 to 2002, pointed out that while there could be “no one single human genome”, his company had indeed “determined a single reference sequence”—albeit an unofficial one (quotes in Bostanci, 2006).

Many of the criticisms of reference genomes we have observed involve some conflation of the ways in which a reference genome can represent or ‘stand-in’ for a species. The idea that the reference genome must be *representative* of the species rather than merely being *a representation* writes cheques that reference genomes often cannot cash. This problem arose

when the basis for the IHGSC conception became untenable, as the extent of functionally-significant genetic variation across humans became apparent. This variation became possible to apprehend and record because of the advent of the reference genome, but undermined the idea that it represented the human species in a universal or metaphysical way. It did not undermine the conception proposed by Myers, in which the reference genome was something more like a type specimen. The appreciation of the extent of genomic variation—and the dissatisfaction with the reference genome occasioned by this growing knowledge and the increasing mismatch between this and the IHGSC’s view—has helped effect a change in the nature of the human reference genome.

As a result, the ontological status of the human reference genome and those for other species such as *S. cerevisiae* has evolved. When the newer reference genome of *S. cerevisiae* was announced in 2014 (Chap. 7), subsequent revisions were supposed to incorporate more variation and better represent the species. In the case of the newer pig reference genome released in 2017 and published in 2020, the authors placed great emphasis on the benefits of the new assembly for finding and exploiting different forms of genomic variation. Developing the reference genome to incorporate more variation was less important to them, though, than it was to human and yeast genomicists. While there has been a general change in the ontological status and modality of reference genomes, with more focus on variability, these may not always be as fully realised for some species or carry the same weight relative to other avenues by which post-reference genomic resources can be developed.

In the case of the human reference genome, having originally been something more like a type specimen (an arbitrary extraction from the diversity of variation found in nature), it has been shifting to become something more like an idealised normal genome, reflecting common non-pathological variants found across populations. This transition constitutes one from the reference genome being an abstraction to becoming more of an idealisation.² What does this mean? As an abstraction, it has been based on the omission of genomic variation through a selective process that depended on multiple choices made throughout all of the stages resulting in the production of a reference sequence. This selectivity has

²There is a rich philosophical literature on abstraction and idealisation. Here, we have deployed some of the senses captured in Cartwright (1989), Godfrey-Smith (2009), Jones (2005), Levy (2018) and Love (2010).

not, though, necessarily been to create a product that is representative or normal in the sense of being only comprised of the most common or non-pathological variants. At the stage of the inception of reference genomes, these were not known, and therefore this was not possible to do. Only with the subsequent apprehension of variation and its functional significance can reference genomes be shaped to take account of—or even incorporate—the common and non-pathological.

Arising from this appreciation of genomic variation, in conjunction with existing ambitions to represent humankind, revisions of the reference genome increasingly tend towards idealisation. It now becomes possible to state that a reference genome is, to some degrees and in some respects, a misrepresentation, as there are now concrete epistemic goals directed towards a specific representational target. This signifies a shift from the dominant epistemic goals of the abstraction phase, which emphasised the contiguity, coverage and quality of assembly—and level of annotation—of the reference genomes. In the idealisation phase, due to the added normative dimensions and the new role that the reference genome is being asked to fulfil, a gap begins to be perceived between the genome itself and the representativeness that it is supposed to embody.

The implication of our transformed picture of the nature and role of the reference genome is not that equality and social justice concerns about the representational scope of genomic resources are invalid. Instead, we would direct such critiques from the reference genome towards the wider webs of reference and observe that such concerns become more salient as one enters deeper into the idealisation phase. Considering the reference genome as a dynamic object that is created and transformed through recursive and iterative processes involving—and sometimes excluding—particular communities of practitioners, is crucial if we are to avoid conflating different ways in which a reference genome can ‘stand-in’ as a representation of a species. As we observe in the next section, the alignments between the aims of genomic research and the concrete processes of idealisation are crucial to effecting translation.

8.3 GENOMICS AND TRANSLATION

The advent of a reference genome is a significant event for any community concerned with the genetics of a particular species. In providing researchers with a comprehensive consensus sequence of the target species, the

reference genome constitutes a resource to which existing and newly-determined genomic data can be related and aligned. It informs the assembly and annotation of new sequences—such as of specific pig breeds or human populations, or different microorganismal strains—and also provides a basis for intra and inter-species comparison.

Particular configurations of pre-reference genomics, and the decisions made in them and in the determination of the reference sequence, affect how readily certain forms of variation can be explored in post-reference genomics. In yeast, there was a pragmatic decision to focus on one particular strain, and this shaped the trajectory of research after the release of the reference genome: participants in the EUROFAN project to functionally annotate the reference sequence were largely drawn from the prior Yeast Genome Sequencing Project. For the human, the gap between the producers of the reference sequence and the medical genetics community led to problems in squaring variation—at least the variation on which medical geneticists had worked before and during the production of the reference genome—with the reference sequence. In pig, although the ‘thin’ compilation of the reference sequence was delegated to the Sanger Institute, significant community continuity happened at the level of the thick sequencing: mapping, assembly and annotation practices. This allowed pig genomicists to appreciate what variation was incorporated in the reference genome, what was missing and what further work needed to be done to characterise different kinds of variation across the species.

The kind of epistemic iteration concerning the development of new genomic resources implies that the characterisation of variation beyond the reference sequence is the central epistemic task of post-reference genomics. Post-reference genomics research involves, in one way or another, the identification, cataloguing, control and use of variation. What variation is being compared, over what time-frame, how it is to be measured, and for what purpose, is up to the researchers involved, who work within various material, theoretical and technical constraints. There are, conceivably, unlimited ways in which comparisons between two (or more) parts, individuals or groups can reveal variation. The particular means by which variation is generated, apprehended, identified, measured, recorded and integrated with other types of variation, conditions (but does not fully determine) the further use that may be made of it.

Sufficiently rich webs of well-connected resources represent different kinds of variation. Key here is the creation of the resources, the processes by which they instantiate particular kinds and ranges of variation, and the data

and material linkages and connections established between them. Capturing extra dimensionalities of data—for instance, through annotation, cataloguing of sequence variants and generating non-genomic biological data—ensures this. So too does apprehending diversity by using the reference sequence, in whole or in part, to characterise specific breeds, strains, populations or even individuals. These practices seed comparisons that enable further functional analysis and the apprehension and detection of variation.

Following on from the points made in Chap. 7 about the development and intersection between the functional and systematic aspects of post-reference genomics, we suggest that translation involves the establishment of means to integrate, link and compare data of these different kinds: those that are associated with phenotypic effects as well as those that pertain to intra- or inter-specific patterns in the sequences. This need not involve a collapsing of distinctions between these modes, but require the alignment and commensuration between resources representing—and derived from—different sources and kinds of data.

In foregrounding alignment in this way, we therefore present a concept of translation that echoes previous social scientific scholarship (Lowe et al., 2020; Sunder Rajan & Leonelli, 2013). This is an interpretation that has more in common with Michel Callon's sociology of translation (Callon, 1986) than with the common use of translation as a policy category concerning the strategy and governance of scientific research. For Callon, achieving translation involves the shaping of a network of actors in a way that structures relations and actions around particular problems and solutions that are posed by one (or more) of the actors. In this way, translation involves “creating convergences and homologies by relating things that were previously different” (Callon, 1980, p. 211; as cited in Wæraas & Nielsen, 2016), be they biological objects themselves or the scientific groups and communities oriented towards them, and their ongoing practices and organisations. These convergences and homologies of biological entities and communities require alignments and commensurations of norms, organisational models and genomic resources.

To adapt Callon's analytical framework into the domain of genomics, we can say that the process of translation consists in defining the epistemic goals (the problems) and determining the means by which these epistemic goals are worked towards (the solutions). In genomics, these processes operate at multiple levels and present different casts and configurations of actors, although there are undoubtedly multiple different overlaps and relations between them.

A main level of operation is the creation of a reference genome, both as a generic object and in specific instances. Generically, the process involves the creation of the category of reference genome (with different designated levels of quality and completeness) by large-scale data infrastructures, such as the RefSeq database of the US National Center for Biotechnology Information, and control over revisions to reference genomes by bodies such as the Genome Reference Consortium. This work creates objects that are commensurate with other forms of genomic (and other omic) data and with other reference sequences (Chap. 1). Yet, before their entrance and commensuration within these centralised infrastructures, genomic data has been produced by different processes involving distinct modes of interaction between the target species and specific communities of genomicists.

Here, we can make sense of some of the different trajectories we have observed for the three species we examined. For the human genome, Callon-style translation was achieved by a small group of actors, primarily at the US Department of Energy, the National Institutes of Health, the Wellcome Trust and some large-scale sequencing centres. They successfully designated the quick generation of a common, accessible reference sequence as the main problem—the epistemic goal of whole-genome sequencing—and so sidelined medical geneticists. An alternative attempt at translation around the sequence produced by Celera provided a more amenable alignment with the interests, practices and norms of the medical geneticists. However, while this enabled some medical geneticists to advance their research, and to produce some genomic resources of use to the wider community (García-Sancho, Leng, et al., 2022), the way that Celera’s data was released—in terms of both access and format—restricted the availability and linkage opportunities around their sequence. Only recently, through initiatives such as ClinGen/ClinVar and the 100,000 Genomes Project, has there been a concerted effort to align large-scale genomics data infrastructures with the interests of, and data produced by, medical geneticists.³

The situation, as we have seen, was quite different in the cases of pig and yeast. There, existing communities working on those organisms achieved translation mostly on their own terms, and this has enabled them

³On the 100,000 Genomes Project, see Jarmo de Vries’s blogpost concerning his ongoing PhD research at: <https://genomicsincontext.wordpress.com/2021/06/11/the-100000-genomes-project-shaping-genomic-medicine-in-the-nhs/> (last accessed 20th December 2022).

to pursue post-reference genome research and alignment with their working world concerns—domains of application such as medicine, agriculture and biotechnology (Agar, 2020)—more-or-less seamlessly, and successfully. For them, the defined epistemic goals shared some commonalities with human genomics, but presented distinct problems that required different solutions to be provided by the reference genome and subsequent resources built using it and relating to it.

In the case of yeast, there needed to be an immediate connection to the experimental practices and aims of the researchers involved, and this meant generating a sufficiently well-annotated reference sequence to enable further exploration through extensive deletions—knockout experiments—and laboratory assays to functionally analyse the genome and its products. This went well beyond the functional annotation that the genome centres initially pursued on the human reference sequence.

For the pig community, the genome simply had to be good enough to enable the selective annotation and further biological explorations of certain regions known to be associated with traits of interest for breeding, developing the pig as an animal model, and furthering the utility of the pig in transplantation biology and xenotransplantation. Additionally, it had to provide the basis for the identification of multitudes of genetic markers such as Single Nucleotide Polymorphisms (SNPs) that would constitute the foundation for new methods of breeding based on the use of these masses of markers. These variants and tools, such as the SNP chips, also furthered the characterisation of the genetic diversity and patterns of distribution of pigs, contributing towards the synergistic relationship between functional and systematic modes that had been a part of pig genomics since the mid-1990s (Chap. 7).

This brings us to another level at which the processes of genomic research operates: concerning the relation of a reference genome to wider webs of reference, and these to forms of biological variation that are pertinent to various working worlds. Alignments and commensurations of the reference genome to various forms of variation enable data, information and interpretations of all kinds to travel through networks of inference and meaning. Working with and beyond reference genomes engenders a greater appreciation of the extent and biological significance of different forms of genomic variation. It also leads to the collection and analysis of data concerning other forms of variation: transcriptomic, metabolomic, all the way to phenotypic, population and community-level. Data concerning these kinds of variation can be linked and related to each other, and to the

web of genomic data. We have seen that this post-reference genomics has informed revisions to reference sequences, and even a shift in the nature of the object of the reference genome. In turn, however, the content of the reference genome conditions what and how new forms of variation can be apprehended and made sense of.

The processes of abstraction of variation involved in the creation of the reference genome, therefore, shape the subsequent idealisation of it and its connections to other reference resources and biological data and materials. The interests of the genomicists that were involved in the production of reference genomes affect their capacity for seeding and influencing the development of subsequent webs of reference. It is within the affordances of the data and materials that result from the historical development of reference genomics that new interconnected nodes can be placed in the abstract variational space that the web of reference ‘explores’. This placement and evolving topology of the web depends on which forms of variation (which new abstractions or idealisations) the communities involved want to generate, to aid the purposes of their research goals and tackling of working world problems. It is easy to see, based on this, that continuity between those actors that successfully seeded and shaped the early development of the web and those actors connected to working world concerns (e.g. in agriculture, biotechnology or medicine) increases the chances of effecting agricultural, biotechnological or medical translation. In other words, Callon-style translation in the production of a reference genome is an important factor in easing or hindering the translation of genomic data and other resources towards addressing practical research problems.

We cannot, of course, consider such webs only in isolation: though they may be furnished with rich internal connections, they undoubtedly also connect to other webs. These connections may be between webs pertaining to different species, but not necessarily, as there may be distinct webs closely associated with particular working worlds. Here, some of the key alignments consist in forming comparative relationships and interoperability between the resources in each of the webs. Again, this is historically conditioned. The extensive development of a comparative inferential architecture between pig and human genomics (Lowe, 2022) aids alignments between webs representing those species. The types and densities of connections and the topology of the ecology of webs depend on socio-historical factors and the nature of the organisms being worked with. On the basis of pan-species projects such as Génolevures, for example, we may expect stronger connections and perhaps fuzzier distinctions between the webs of reference of different species of yeast.

Finally, there is the more general level of the overall infrastructure and norms of genomics. At this level, the actors who successfully achieved and built on their Callon-style translations at other levels may have more subordinate roles or at least less dominant ones. This level, consisting of the data infrastructures and their associated rules and norms, institutions, funding and publication policies, and even a certain vision of what genomics is and should be, has been strongly bent in the direction of the problems and solutions presented by those core actors that directed the production of the human reference sequence. Because of this, some species such as yeast with the resources and disposition of a model organism community, as well as a history of genomics that precedes the completion of the human reference genome, may exhibit more independence than pig genomics, which conducted its sequencing afterwards and always had strong connections with the mainstream of human genomics. The pig genomics community, though, has been able to shape genomics in a way congenial to the aims and interests of the genomicists comprising it. Their existing working world ties to the breeding industry have provided a means for the recapitulation of the pre-genomic norms of animal genetics into farm animal genomics.

Sociological translations are involved at all these levels, each of which have been shaped by the ways that different communities of genomicists have been formed and their attempts—and differential success—at effecting the translation of their interests. All these levels and factors have conditioned the development of reference genomes and subsequent webs of reference. These genomes and webs of reference, in turn, affect how the tools for the further characterisation of data concerning variation can align with working world problems, be it medical genetics, livestock breeding or the investigation of a model organism. As well as furnishing the socio-historical conditions affecting the chances of successful scientific translation, these processes also shape what medical and agricultural applications are considered doable or desirable.

Further characterising webs of reference and the nature of post-reference genomics is a vital task. It will require working across methods and disciplines, combining more conventional historical and philosophical inquiry alongside qualitative and quantitative methods in the social sciences. It also will require an engagement with, and sensitivity to, the concrete paths of research developed across different domains of species and working world orientations. We close this concluding chapter with reflections on some methodological aspects that future research should take account of, concerning the periodisation and demarcation of genomic research.

8.4 PERIODISATION, MULTISPECIES APPROACHES AND COMMUNITIES AS HISTORICAL ACTORS OF GENOMICS

One of the main arguments of the book has been to distinguish between a historical periodisation that strictly identifies an age of genomics (roughly 1990 to 2003, with post-genomics succeeding it) and our narrative in which genomics is an ongoing enterprise, albeit featuring distinctive shifts in the organisation and nature of the endeavour following the production of a reference genome. Our interpretation takes fuller account of the differential historical trajectories of genomics concerning different species and the communities that worked on them. It also stresses the fact that the practices and outputs of genomics continue into the so-called ‘post-genomic’ era. Furthermore, in certain communities and genomic enterprises concerning particular species, constitutive features that scholars have attributed to post-genomics (e.g., in Richardson & Stevens, 2015) were also present in genomic research.

Genomics is not a discrete—nor complete—phase of scientific endeavour. It is continually transformed and enters into new combinations and relations with other data being generated and handled in particular ways. Our notion of post-reference genomics captures this, but also encapsulates the situatedness and historicity of particular strands of post-reference genomics that deal with specific objects, such as species or groups of related species. While post-reference genomics represents a category, it can be manifested in distinct ways by different communities, in different time periods and with differing consequences. We can therefore observe diverse historical trajectories, other than those of the canonical periodisation of genomics centred on the ‘completion’ of the human reference genome and the alleged start of a new post-genomic era (Fig. 8.1).

As we have shown in Chaps. 3 and 4, even for *H. sapiens* there were a plethora of initiatives that, while directed to the human genome as an object, did not pursue the production of a full reference sequence. Because of this, these initiatives did not adopt the genome centre model or industrial forms of organisation aimed at the rapid production of a whole-genome sequence. Rather than deploying large-scale approaches, these initiatives sought to map and sequence targeted areas of the genome. The distance between the communities of medical and human geneticists that undertook these initiatives and the producers of the reference genome created a perceived ‘translational gap’ around the exploitation of the clinical and scientific potential of the full sequence. In other words, the distinct

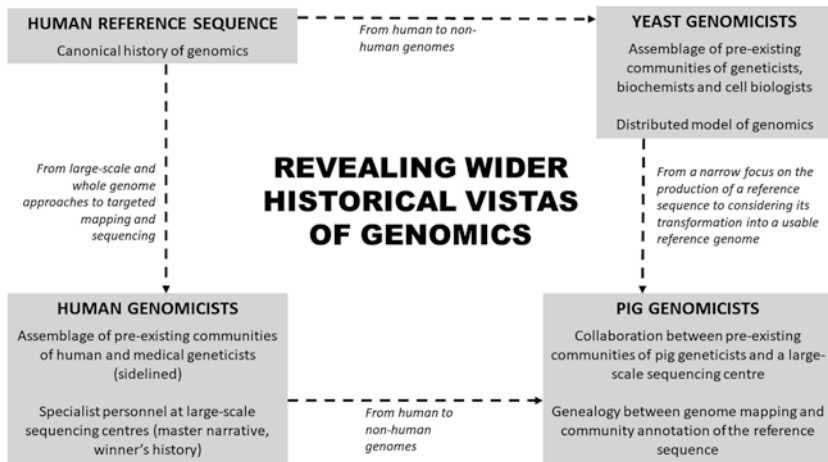


Fig. 8.1 A diagrammatic representation of how an emphasis on the interactions between different communities and their target genomes expands the historical vistas of genomics. Dotted lines represent our historiographical de-centring from the production of the human reference sequence. Below each community of genomics, we outline how their trajectory diverges from the canonical history of genomics. Elaborated by both authors. For a larger version of this figure that can be zoomed in and out, see https://www.pure.ed.ac.uk/ws/portalfiles/portal/290406893/Fig_8_1_increased_final.pdf

historicities and motivations of two different communities of genomicists—human and medical geneticists, on the one hand, and more specialised operatives at genome centres, on the other—created a disjunction between the reference sequence produced by just one of them, but that was intended for use by the other.

If we shift from the human to non-human species, we observe that while yeast and pig genomics sought the production of a full reference sequence, their historical trajectories differ from the canonical one. For yeast, a long-established, tight-knit community working on a specific strain of *S. cerevisiae* decisively contributed to the production of the reference sequence in what we called the distributed model of genomics, as opposed to the concentrated determination of the human reference genome at specialised sequencing centres (Chap. 2). Pig genomics squares with the canonical trajectory if we consider the ‘thin’ production of the reference sequence; after all, this endeavour was modelled on the plans and methods of the international consortium that produced the

equivalent sequence for *H. sapiens*, and it was largely undertaken by the Sanger Institute. Yet if we consider the ‘thicker’ practices that were involved in making this sequence a robust reference resource, other genealogies become apparent and challenge the rigid periodisation of genomics and post-genomics. For instance, the agriculturally-inclined geneticists and immunogeneticists involved in the prior mapping of the pig genome were crucial in its community annotation, which required collaboration between the Sanger Institute and those long-established pig genomicists (Chaps. 5 and 6).

Although our perspective de-centres the human reference sequence as the paradigmatic—even definitional—instantiation of genomics, it does not necessarily remove it from an important role in the shaping of the history of genomics more broadly. Instead, it calls attention to examining the concrete ways in which this reference sequence, and more specifically the idea of one Human Genome Project that produced it, generated a gravitational attraction around the version of genomics it embodied. As we have shown throughout the book, this centripetal force was associated with broader socio-political processes and, crucially, established retrospectively in the accounts of James Watson and other prominent participants. The master narrative of genomics, centred on the idea of a single and successful Human Genome Project, was—and is still—influential because of its alignment with other influential historical forces, not because it represents an intrinsically superior or dominant way of conducting science.

There is a tension implicit in our de-centring and alternative periodisation of genomics. Through identifying the advent of a reference genome as an inflection point, rather than a transition to a wholly new post-genomic endeavour, we appear to suggest that the structure of the history of genomics differs according to the species. After all, while yeast entered our proposed post-reference genomic period in 1996, the human did not do so until 2003, and the pig until 2011.

This historiographical transition from a human-centred periodisation towards one based on species-specific designations of pre-reference, reference and post-reference genomic phases constitutes an advance in appreciating the heterogeneities and continuities we have observed in this book. It, however, still constitutes an incomplete and patchy picture. This is because, in spite of the distinct periodisations for each species, the overall development of genomics—its infrastructures, norms, data, materials, methods and techniques—possesses its own rhythm and historicity. These may have developed out of one or a few distinct initiatives—such as

Ensembl and the Human and Vertebrate Analysis and Annotation (HAVANA) group being born out of the human reference genome sequencing programme—but once created have had a life, development and impact beyond them. It matters for understanding some of the differences between the histories of pig and yeast genomics that an existing sequencing, assembly and annotation infrastructure was in place at the Sanger Institute for pig genomics but not for yeast, for example. And in turn, it is consequential that the particular way in which pig genomics developed affected the way that HAVANA, in particular, changed in the post-human reference genome era.

Our approach to the history of genomics has enabled us to identify this relationship between more global and local repertoires, processes and configurations. As well as de-centring from the illusion that one model—the Human Genome Project—is generalisable, it has helped us to unpick the commensuration work of administrative agencies and large-scale infrastructures, such as the RefSeq database. It has also enabled us to reveal the historical trajectories that give the products of genomics research different affordances and limitations. The usefulness of our approach is not restricted to being merely comparative; it also enables connections to be identified. It remains an open question how best to harmonise—or, at least, operationalise—the always conflicting tension between histories that are strongly species-specific and those that concern the more general development of genomics as an infrastructural and data-centred endeavour. We have, we hope, now opened up the space for such questions to be asked and explored.

REFERENCES

- Agar, J. (2020). What is science for? The Lighthill report on artificial intelligence reinterpreted. *The British Journal for the History of Science*, 53(3), 289–310.
- Ankeny, R. A. (2000). Fashioning descriptive models in biology: Of worms and wiring diagrams. *Philosophy of Science*, 67(Proceedings), S260–S272.
- Ankeny, R. A. (2007). Wormy logic: Model organisms as case-based reasoning. In A. N. H. Creager, E. Lunbeck, & M. N. Wise (Eds.), *Science without laws: Model systems, cases, exemplary narratives* (pp. 46–58). Duke University Press.
- Ankeny, R. A., & Leonelli, S. (2011). What's so special about model organisms? *Studies in History and Philosophy of Science Part A*, 42(2), 313–323.
- Ballouz, S., Dobin, A., & Gillis, J. A. (2019). Is it time to change the reference genome? *Genome Biology*, 20, 159.

- Barnes, B., & Dupré, J. (2008). *Genomes and what to make of them*. The University of Chicago Press.
- Bolker, J. A. (2017). Animal models in translational research: Rosetta Stone or stumbling block? *BioEssays*, 39(12), 1700089.
- Bostanci, A. (2006). Two drafts, one genome? Human diversity and human genome research. *Science as Culture*, 15(3), 183–198.
- Callon, M. (1980). Struggles and negotiations to define what is problematic and what is not; the socio-logic of translation. In K. D. Knorr, R. Krohn, & R. Whitley (Eds.), *The social process of scientific investigation* (pp. 197–219). D. Reidel.
- Callon, M. (1986). Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay. In J. Law (Ed.), *Power, action and belief: New sociology of knowledge?* (pp. 196–229). Routledge & Kegan Paul.
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Clarendon Press.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford University Press.
- de Chadarevian, S. (2020a). Normalization and the search for variation in the human genome. *Historical Studies in the Natural Sciences*, 50(5), 578–595.
- de Chadarevian, S. (2020b). *Heredity under the microscope: Chromosomes and the study of the human genome*. The University of Chicago Press.
- Gannett, L. (2003). The normal genome in twentieth-century evolutionary thought. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 34, 143–185.
- Gannett, L. (2019). The Human Genome Project. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition). Retrieved December 20, 2022, from <https://plato.stanford.edu/archives/win2019/entries/human-genome>
- García-Sancho, M., Leng, R., Viry, G., Wong, M., Vermeulen, N., & Lowe, J. W. E. (2022). The Human Genome Project as a singular episode in the history of genomics. *Historical Studies in the Natural Sciences*, 52(3), 320–360.
- Garner, J. P. (2014). The significance of meaning: Why do over 90% of behavioral neuroscience results fail to translate to humans, and what can we do to fix it? *ILAR Journal*, 55(3), 438–456.
- Godfrey-Smith, P. (2009). Abstractions, idealizations, and evolutionary biology. In A. Barberousse, M. Morange, & T. Pradeu (Eds.), *Mapping the future of biology: Evolving concepts and theories* (pp. 47–56). Springer.
- Holmes, T. (2017). The wild type as concept and in experimental practice: A history of its role in classical genetics and evolutionary theory. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 63, 15–27.
- Jones, M. R. (2005). Idealization and abstraction: A framework. In M. R. Jones, N. Cartwright, & N. (Eds.), *Idealization XII: Correcting the model; idealisation and abstraction in the sciences* (pp. 173–218). Rodopi.

- Leonelli, S. (2016). *Data-centric biology: A philosophical study*. The University of Chicago Press.
- Leonelli, S., & Ankeny, R. A. (2013). What makes a model organism? *Endeavour*, 37(4), 209–212.
- Levy, A. (2018). Idealization and abstraction: Refining the distinction. *Synthese*, 198, 5855–5872.
- Love, A. C. (2010). Idealization in evolutionary developmental investigation: A tension between phenotypic plasticity and normal stages. *Philosophical Transactions of the Royal Society B*, 365, 679–690.
- Lowe, J. W. E. (2016). Normal development and experimental embryology: Edmund Beecher Wilson and *Amphioxus*. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 57, 44–59.
- Lowe, J. W. E. (2022). Humanising and dehumanising pigs in genomic and transplantation research. *History and Philosophy of the Life Sciences*, 44, 66.
- Lowe, J. W. E., Leonelli, S., & Davies, G. (2020). Training to translate: Understanding and informing translational animal research in pre-clinical pharmacology. *Tecnoscienza*, 10(2), 5–30.
- McKusick, V. A. (1989). Mapping and sequencing the human genome. *The New England Journal of Medicine*, 320, 910–915.
- Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bizakadze, A. V., & Mikheenko, A. (2022). The complete sequence of a human genome. *Science*, 376, 44–53.
- Olson, M. V. (2011). What does a “normal” human genome look like? *Science*, 331(6019), 872.
- Richardson, S. S., & Stevens, H. (Eds.) (2015). *Postgenomics: Perspectives on biology after the genome*. Duke University Press.
- Rosenfeld, J. A., Mason, C. E., & Smith, T. M. (2012). Limitations of the human reference genome for personalized genomics. *PLoS ONE*, 7(7), e40294.
- Sunder Rajan, K., & Leonelli, S. (2013). Introduction: Biomedical trans-actions, postgenomics, and knowledge/value. *Public Culture*, 25(3), 463–475.
- Tauber, A. I., & Sarkar, S. (1992). The Human Genome Project: Has blind reductionism gone too far? *Perspectives in Biology and Medicine*, 35(2), 220–235.
- Wæraas, A., & Nielsen, J. A. (2016). Translation theory ‘translated’: Three perspectives on translation in organizational research. *International Journal of Management Reviews*, 18(3), 236–270.
- Witteveen, J. (2016). Suppressing synonymy with a homonym: The emergence of the nomenclatural type concept in nineteenth century natural history. *Journal of the History of Biology*, 49, 135–189.
- Witteveen, J. (2018). Typological thinking: Then and now. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 330, 123–131.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

