Abstract

One of the defining features of the classical gene was its *position* (a band in the chromosome). In molecular genetics, positions are defined instead as nucleotide numbers and there is no clear correspondence with its classical counterpart. However, the classical gene position did not simply disappear with the development of the molecular approach, but survived in the lab associated to different genetic practices. The survival of classical gene position would illustrate Waters' view about the practical persistence of the genetic approach beyond reductionism and anti-reductionist claims. We show instead that at the level of laboratory practices there are also reductive processes, operating through the rise and fall of different techniques. Molecular markers made the concept of classical gene position practically dispensable, leading us to rethink whether it had any causal role or was just a mere heuristic.

Keywords: gene position, reductionism, molecular markers, genetic approach

1. Reductionism in the practice turn

The practical turn in philosophy of science has led to a reconsideration of many classical concepts in the field. The debate on reduction during the last half of 20th century hinged around theories, whether some of them could be somehow inferred from some other theories, in a way that captured the progress of science. Attractive as it was, there have been very few consensually successful reductions and this lack of results made anti-reductionist pluralism the mainstream view among philosophers of science. Ken Waters has recently restated in practice terms this debate between reductionists and anti-reductionists, drawing on an analysis of genetics throughout the 20th century.

The reducibility of classical to molecular genetics has been a matter of controversy for decades in the philosophy of biology. The classical definition of gene has four principal elements (Weber 2004): position, mutation, recombination and function. In classical genetics, with the rise of cytogenetic techniques, genes became the unit for function, recombination and mutation while keeping a position on the chromosomes. Genes functioned as causes for the phenotype, and were assumed to be the target of mutations causing variations in those phenotypes[1].

---

[1] As a matter of fact, genes were the elements of recombination as this process would shuffle the linkage between alleles (Pierce 2013).

The development of molecular genetics presented genes as the class of DNA sequences that determine the linear sequences of amino acids in a protein (Weber, 2004)[2]. This new gene concept was completely defined in molecular concepts yielding the description of the central dogma of molecular biology. Molecular genetics also led to a reinterpretation of the four fundamental classical features in terms of DNA. We now consider that the unit of mutation is the single nucleotide while the unit of recombination is a restricted number of nucleotides, as small as two but probably expanding to twelve (Benzer 1955) (Chakraborty et al. 2018). However, it is still debated whether function and position have any clear DNA translation[3].

There are other coincidences between other classical and molecular gene features, suggesting for some a potential overlap between the two concepts (Baetu 2010). However, the correspondence is far from complete and this has been *prima facie* an argument for the anti-reductive pluralist. Waters suggests instead to consider this lack of correspondence from the standpoint of genetic practices. Focusing on how geneticists work in their laboratories, Waters argues that geneticist often retool classical approaches into molecular terms. Inasmuch as this retooling leads to a better grasp of mutations and their effects, any combination of classical and molecular principles should be acceptable, whatever the theoretical or ontological implications.

We want to suggest instead that even within the realm of practice there is still room for reduction, now understood as conceptual dispensability. Different theoretical approaches do actually coexist in genetics laboratories, but only to the extent that they help practitioners to achieve certain epistemic benchmarks. If this same goal suddenly can be accomplished with new techniques in which some of the older approaches play no role, we will say that these older approaches have become practically reduced.

This is the basic intuition that we want to substantiate through an analysis of classical and molecular concepts of gene positions, showing how they have coexisted in the laboratory until recently, when molecular markers made classical positions practically dispensable. Since our argument is a development of Waters, we will start, in section 2, presenting his own views

---

[2] This is a broad, consensual, definition of the molecular gene: for further details, see (Waters 1994) and (Griffiths and Stotz 2013).

[3] A reviewer objects that, at no point in our argument, "are *bona fide* molecular concepts of the gene introduced or discussed (e.g., open reading frames, transcription units, annotated genes)". These are concepts that allow geneticists to reconstruct some classical gene properties in molecular terms. But we are going to focus on classical gene position, a concept for which no such reconstruction has been achieved. Therefore, those *bona fide* molecular concepts will not feature in our analysis.

about reductive retooling in genetics as a third way between reductionist and antireductionist claims. In section 3, we will illustrate in detail the reductive retooling of classical gene positions in molecular terms. Our test case will be the investigation of the genes responsible for a particular trait in a model organism (the white eye of *Drosophila*). Here we will see different implementations of what Waters calls the *genetic approach*, a general heuristic to investigate mutations in classical and molecular genetics. Although conceptually different, classical and molecular gene positions were used simultaneously in order to identify the precise region of the X chromosome where the trait originated.

Unlike Waters, we are going to defend next that we should not only focus on processes, but on outputs as well. Practice is not only based on research strategies, but also on epistemic benchmarks arising from those same practices. In section 4, we will discuss how molecular geneticists consider accomplished the identification of the genes responsible for a particular trait when they put together a statistical and a mechanistic description of the mutation under study. Geneticists need, on the one hand, a statistically significant association between a position in the genome and a phenotype and, on the other hand, a full analysis of the mutation mechanism. Although a fully accomplished description of a mutation is rare, it provides an actual epistemic benchmark in molecular genetics (e.g., for database curators recording their findings).

In section 5, we show how this epistemic benchmark allows us to establish the dispensability of classical concepts. Whereas for several decades classical and molecular positions coexisted in the implementation of the genetic approach, the development of molecular markers made classical position dispensable: a fully accomplished explanation of the development of a phenotypic trait could be achieved without invoking classical concepts. The development of molecular markers makes classical gene positions dispensable. In section 6, we will present such dispensability as a form of reduction, as opposed to traditional forms of epistemic or ontological reduction. Following Waters' inspiration, we will call it a *practical* reduction. Section 7 concludes.

2. Waters' reductive retooling

(Waters 2008) distinguishes between *theoretical reductionism* and *layer-cake antireductionism*. The former would require the reduction of the classical theory of genetics to a theory about molecular processes involving DNA. Such theoretical reduction would imply that all (theoretical) concepts used and phenomena tackled in classical genetics are to be reinterpreted in molecular terms (basically, using our knowledge of nucleic acids and other

3

molecular factors). *Layer-cake antireductionism* argues that classical and molecular genetics explain different phenomena. In this case, classical genetics would have already addressed the core phenomena and the molecular theory would only add explanations for 'peripheral' principles of classical genetics.

According to Waters, both reductionists and anti-reductionists are equally deluded about the actual practices of genetics. Both approaches focus on theories and explanations, disregarding the research methods that geneticists use. For Waters, these approaches represent *fundamentalist* views: they are traditional analyses of theories exclusively aimed at describing the fundamental causal relationships in a specific world domain (Waters 2014). In the case of genetics, (Waters 2008) states that "these views focus on how geneticists explain or try to explain phenomena, not on how they manipulate or investigate phenomena". Waters even considers that (Wimsatt 1976) antireductionist account of *heuristics* in genetics is solely focused on (theoretical) explanation, and thus it is a fundamentalist view.

Waters suggests instead an alternative *toolbox* view focused on practices of theorizing (Waters 2014). The objective of these practices is to construct causal models that explain aspects of the processes in a domain. As these practices change, they provoke a retooling that may yield new models. Old and new models may co-exist as they explain better different aspects of their domain. For Waters, there is no reason to assume that all these models will be once "subsumed under a single set of fundamental principles".

Following (Waters 2008) again, the distinctive trait of these practices is how geneticists recast ideas of classical genetics in terms of molecules, so that they can build upon the investigative strategies of the former using molecular conceptual tools. This investigative strategy of genetics is what he calls the "genetic approach", which would proceed in three steps:

> (a) find naturally occurring or artificially produced mutants that exhibit a difference relevant to some biological process of interest, (b) carry out genetic analyses of the mutants, and (c) recombine the mutants to learn more about the process of interest.

For Waters, the genetic approach allows the use of genetics to investigate a broad range of phenomena, many of which are quite distant from the processes directly involving DNA (e.g., the growth of axons in the nervous system of nematodes). The genetic approach allows the manipulation, control and understanding of these processes, and the exportation of genetic methods. This export takes the form of a *reductive retooling* (Waters 2014). Instead of reducing theoretical principles, genetics makes progress recasting the classical genetic approach in terms

of a different procedural know-how -e.g., the use of molecular-based methods to identify genes rather than transmission-based methods. In Waters' view of genetics, inter-theoretical reductions would play no role, since the practices under analysis are indeed expansive: genetics+, for Waters, is about biologists using the genetic approach beyond the traditional realm of genetics, and there is no point for those practitioners in reducing their theories to molecular genetics.[4]

In our view, Waters can only claim that reductionists and anti-reductionist are misled about genetic practices, because he focuses on practices outside genetics *sensu stricto* where there is little room for reduction -the underlying theories are perhaps too distant. But if we focus instead on practices within the traditional realm of genetics, like the explanation of phenotypic variation, we will see that the implementation of that same genetic approach has reductive consequences. More precisely, the laboratory practices of molecular genetics trigger the same sort of eliminative processes conceptualized so far as reductions, in which theoretical considerations still play a bigger role than Waters seems willing to admit.

In order to make our case, let us first illustrate how the transition from classical to molecular genetics can be interpreted in terms of a retooled genetic approach. Our running example will be one of the first traits to be studied in genetics, the wrinkled-seed character of pea plants described by Mendel –a more detailed illustration will follow in section 3. In this case, the identification of naturally occurring mutants (in his garden) and the genetic analyses of the mutants (the experimental crosses) were achieved by Mendel himself, still in classical terms – step 1. A further step was taken by (Bhattacharyya et al. 1990) when they identified a transposon-like insertion as being the actual mutation in the DNA sequence behind the effects, using molecular techniques such as DNA sequencing. In this same paper we also find the third step in the genetic approach in the form of recombining experiments articulated, again, in molecular terms.

This reductive retooling just replaces classical for molecular models and techniques, leaving intact the guiding heuristic (the three steps genetic approach). It would be a third way between *theoretical reductionism* and *layer-cake antireductionism*, explaining the advancement of genetics in terms of a succession of models driven by a general practical guidance for research on mutants.

---

[4] We are grateful to Ken Waters and an anonymous reviewer for helping us to clarify this point.

We are now going to dig deeper into what is precisely *reductive* in a reductive retooling focusing on classical and genetic approaches to gene position. Let us start, in the following section, illustrating both approaches at work in the reductive retooling of gene position that took place when geneticists investigated the genes responsible for the *Drosophila* white eye phenotype.

3. The molecular retooling of classical gene position

Drosophila is a model organism: a small fly with red eyes. The white eye was a natural mutation that, according to Waters' template, prompted classical geneticists to start their analysis and reproduce the mutants leading to the discovery of a chromosomic locus (a gene position) responsible for the phenotype.

Among classical geneticists, (Morgan 1910) and (Bridges 1914) are perhaps the most influential references on the role of chromosomes in the inheritance of traits. They gathered enough data to detect a statistical association between a locus in the X chromosome and the white phenotype in two-generation experimental crosses. Their (classical) genetic approach led to the identification of a rough gene position (a chromosomic band) from which the white phenotype originated. Further crossing allowed for a more detailed mapping of the locus in the following decades –e.g., (Green 1959), after cytogenetic banding techniques were improved[5].

The molecular retooling of the (classical) genetic approach led to a different understanding of the white eye locus. Instead of focusing on segregation in experimental crosses, geneticists proceeded to inactivate various elements at the molecular level (proteins, RNA) to study their phenotypic effects. (O'Hare et al. 1984) targeted proteins involved in the pigment synthesis and obtained white-eyed *Drosophila*. (Levis et al. 1984) worked on targeted mutations caused by the insertion of transposable elements and explored the effects of deactivating the gene at the RNA level, seeking structural or expression deviations in the messenger molecule. The study partially succeeded in showing that inactivation (in this case caused by absent or truncated RNA) explained the effects attributed to the white eye locus in classical genetics. (Gehring et al. 1984) incorporated a fully functional allele into an organism

---

[5] We should bear in mind that cytogenetic techniques and experimental crossing provide different information on gene positions. While cytogenetics identifies chromosomes bands, experimental crosses inform about different genes being located nearby, in the same chromosomic region. Even though cytogenetic techniques dealt directly with the chromosome, classical genetics used experimental crosses to identify gene positions in the newly constructed linkage maps. As we will see later in section 5, the information provided by experimental crosses got into molecular genetics through molecular markers.

with defective white copies, allowing the normal synthesis of the pigment. The authors used a vector carrying a red-eyed haplotype and, since they succeeded in obtaining a white-eye phenotype, they claimed that the DNA fragment in the vector contained the full sequence of the locus.

The progress of experimental practices in genetics is, in this sense, a reductive retooling of a classical approach into molecular terms, in which classical and molecular elements go hand in hand to facilitate the completion of the experiment. In our view, gene position provides an excellent illustration of this retooling. Position in the classical approach refers to a location on the chromosome, which was usually identified through cytogenetic techniques (typically, chromosome banding techniques). In molecular genetics, position refers to a specific nucleotide location in the organized genomic sequence. As we will see below, there is no clear correspondence between the two. In terms of DNA sequences, classical positions represent extremely long stretches of nucleotides. However, knowing these classical positions is extremely useful for molecular geneticists in order to deploy Water's genetic approach. Classical positions allow them to focus on specific sites of the genome to further investigate mutations and their effects through molecular means. In the case of the white eye locus in *Drosophila*, once located in a specific region of the X chromosome, the molecular retooling allowed geneticists to use recombinant DNA techniques to narrow the position and finally establish the complete sequence of the gene.

In our view, the study of *Drosophila*'s white eye nicely illustrates Water's third way: reductionism or anti-reductionism play no role in the actual research practices leading to the identification of the responsible genes. Classical and molecular approaches coexist and geneticists alternate between classical and molecular concepts of gene position in order to implement Waters' genetic approach and narrow down the search for the white eye locus. However, for our argument it is also crucial to consider the difference in output of both approaches. Whereas for classical geneticists the expected outcome was a statistically significant association between a position in the chromosome and a phenotype, their molecular peers aim at obtaining, in addition, a functional description of the mutation mechanisms involved plus the relevant biochemical analyses at all the significant molecular levels. As to the mutation mechanisms, the molecular approach is considered accomplished if geneticists are able to generate a genetically modified organism (unicellular or multicellular). This process confirms the effects of the putative polymorphism, if the genetically modified organism (usually belonging to the same species) replicates the phenotype under study. As to the companion

biochemical analyses, the list is still expanding: currently it includes DNA, RNA, proteins and, most recently, epigenetic modifications. As we are going to argue next, this explanatory pattern (statistical description + full mechanistic description) is central to the actual practice of genetics. The coexistence of classical and molecular concepts is epistemically justified to the extent that both contribute to achieve this output. As we will argue in section 5, once this same output is achieved without classical concepts, a practical reduction will be taking place.

4. Causal explanation in genetics

The first step in our argument is to identify a practical epistemic benchmark for successful explanation of phenotypic traits in molecular genetics. Although the combination of mechanistic and statistical evidence has often been defended as a methodological paradigm - see, for instance, (Russo and Williamson 2007) for the biomedical sciences-, we want our analysis to remain at the level of genetic practices *à la* Waters. If we focus on the curatorial practices of genetic databases, we find implicit *evidence rankings* that suggest that the explanation of the *Drosophila* traits discussed above is indeed an accomplished explanatory combination of statistical and mechanistic evidence.

The Online Mendelian Inheritance in Man (OMIM) is a "comprehensive, authoritative compendium of human genes and genetic phenotypes". Widely used by researchers and clinicians, OMIM targets the connection between phenotype and genotype. Each entry refers to a gene-phenotype relation, containing structured summaries of the relevant information based on expert reviews of the biomedical literature. Each entry is classified in different categories according to the reliability of the association (Amberger et al. 2018). This reliability is often assessed in terms of the evidence contributed. All entries include putative associations between genomic regions (sometimes so narrow that they are, in fact, single genes) to phenotypes. In other words, statistical evidence is the initial requirementff to assess putative relationships between genotypes and phenotypes. When the entry contributes, in addition, mechanistic evidence, it is usually filed in the top category ("Phenotype description, molecular basis known"). Geneticists consider that, for entries in this category, the nature of the relationship between genotype and phenotype is considered well established. The same principles are also implemented in the Online Mendelian Inheritance in Animals (OMIA) database (Lenffer et al. 2006)).

Nonetheless, it is rare to find an accomplished molecular output including both the statistical association and the functional description of the mutation mechanisms. In fact, it has

been achieved mostly in model organisms (in which a variety of molecular tools are available) and for quite simple traits. A well-known example is the seminal analysis of the locus *Extension* (a major gene affecting fur color in animals). This locus harbours the *MC1R* gene, a master regulator of fur coloration and a central piece in the biochemical network of melanin synthesis. (Searle 1968) identified *Extension* and its alleles in mice using experimental crosses. (Falconer 1962) was able to first associate the phenotype with a position in the murine genome, thus gathering most of the information provided by classical genetic approaches. It took decades to identify *MC1R* as the gene responsible for the dominant and recessive phenotypes of the *E* locus in mice (Robbins et al. 1993). This accomplished molecular outcome includes a fine mapping of the locus on the murine chromosome 8 using RFLP molecular markers. It also features a characterization of all the gene variants at the DNA and the RNA level, plus the assessment of the protein function. It therefore includes all the molecular levels (DNA, RNA and protein) and the actual changes that lead to the phenotype (activation or inactivation of a protein receptor). All the relevant biochemical mechanisms have been identified. Moreover, the effects of the polymorphism were confirmed in genetically modified pigment cells. All in all, for most genetics practitioners, this amounts to a full explanation of the effects of the gene *MC1R* and the mechanisms involved in the control of coat color.

This causal explanation combines mechanistic and statistical evidence. In genetics, it is common and even paradigmatic to require both types of evidence to prove that a mutation of the DNA is causing a particular genotype. On its own, evidence about mechanisms is rarely conclusive about the link between a mutation and a phenotype. On the one hand, there are close polymorphisms on the same DNA region (*linkage*), all potential triggers of a particular phenotypic effect. On the other hand, polymorphisms may have several phenotypic effects (including no effects at all) and, in addition, there is no complete knowledge of all the putatively altered biochemical pathways and the influences of genetic contexts (all the other genetic variants of that particular organism). Thus, even if we have a reliable mechanism to relate a mutation and a phenotype, both have to be found together in more than one organism, making statistical evidence indispensable for the analysis. In current molecular genetics, the statistical evidences come mainly from association studies (mainly GWAS) were phenotypes are associated to chromosomic locations. These studies show that specific sequences on the DNA are shared between individuals displaying the same particular phenotype, and are performed by using several hundred molecular markers.

Although the degree of manipulation that molecular mechanisms allowed was obviously beyond the reach of classical geneticists, they nonetheless tried rough forms of chromosome manipulation (e.g. irradiation). As we already mentioned, in classical genetics, candidate mechanisms were roughly identified through the analyses of phenotypes in experimental crosses. The phenotypic frequencies would then provide information about the inheritance pattern of the trait, so that in some cases it could be linked to chromosomic bands. Once it became clear to classical geneticists that mutations could modify the phenotype, describing the mechanisms leading to such a change became one of the central goals of genetics. Yet, classical geneticists focused mostly on describing the direct consequences of random mutations, analyzing their putative effects with the information they had (chemical modifications, chromosome, etc.), with a rudimentary understanding of these mechanisms.

The first molecular experiments were focused on the elimination or the addition of a chromosomic region (previously associated to a phenotype) that would be assayed to replicate the phenotype. With the development of molecular genetics, it became possible to target specific locations on the DNA sequence, with or without alteration of the DNA sequence - including, as we have seen in the hair color example, RNA expression, protein function and even transgenesis and genomic modification. These assays are used to infer the interactions of the elements involved in phenotypic expression, and thus provide mechanistic evidence at the molecular level.

From a molecular standpoint, for instance, it is not very significant to know the (classical) position a new aminoacid will have in a protein due to a mutation on the DNA. Molecular geneticists have discovered that these mechanisms may involve several levels of complexity (from protein function to complicated metabolic interactions), even if every mutation starts with an apparently small modification on the DNA. And they can actually predict conformation changes of the protein. The complexity of classical and molecular mechanisms is of a very different magnitude.

This combination of statistical and mechanistic evidence accounting for the full descriptions of the mutation's effects and its role on the phenotype is indeed rare. As evolutionary biologist Hopi Hoekstra points out, "identifying the precise causal mutations and the mechanisms by which they effect phenotypic variation is a cumbersome task that has only been undertaken by a few intrepid researchers for relatively few loci" (Hoekstra 2014). An objection we should fend off is why taking as paradigmatic something that is so rarely achieved.

If our benchmark are the scientific practices of geneticists, why focusing on such an exceptional output? Our answer is that these full descriptions are exceptional because they are very difficult to obtain. Yet, geneticists do not give up and a great deal of their actual practice is about overcoming the obstacles and achieve a full description.

Let us first discuss the difficulties of identifying a candidate mutation and describe the relevant mechanisms. As to the statistical analysis, there are often many candidate mutations for a given phenotypic effect. We may ascribe this effect to one particular mutation through the analysis of the overrepresented alleles in organisms (of the same species/population) sharing the same phenotype. But it often happens that all the genetic polymorphisms in a given DNA region behave the same and have equally overrepresented alleles –as we already mentioned, this is due to DNA linkage, on which more below in section 5. Under these circumstances, it is very difficult to identify the specific candidate mutation that will allow geneticists to track the relevant mechanisms leading to a particular phenotype.

But even when statistical analysis succeeds, the functional description of the mutation mechanisms is still a very complicated task. On the one hand some mutations are not located on genes and sometimes they are, but we know little about those genes. Some other mutations may have regulatory effects without changing the structure of the protein or with very mild phenotypic effects. Even in those cases in which there is a protein change, these changes may occur at very particular times or tissues in which they are difficult to track. Finally, the process of obtaining a viable genetically modified organism is extremely expensive and often difficult, if not impossible.

Achieving a complete description with the *MC1R* gene was feasible for a number of particular reasons. On the one hand, the coded protein is a relatively small membrane receptor that belongs to the fully described class of G- coupled protein receptors. It was possible to locate the candidate mutation in the correct functional domain. On the other hand, there is a good and very simple model organism (cultured pigment cells) to study pigment synthesis. It was relatively easy to create a genetically modified version of these cells to generate the desired effect –since colour followed a Mendelian pattern of inheritance.

Yet, most genes are not so easy to analyze at the molecular level. For instance, there is a region in the chromosome 4 of pigs that has been linked to fat deposition since 1994 (Andersson et al. 1994). So far, the actual mutation has not been described, although there have been about 20 papers trying to refine the mapping of DNA sequences associated to the fat

deposition phenotype. The list of obstacles is long: the length of the region, non-Mendelian inheritance, etc. Nonetheless, the analysis of candidate mutations goes on and the relevant genetic databases (i.e, OMIM and OMIA) keep gathering data on potential model organisms and molecular mechanisms.

Summing up, difficult as it may be to achieve a full explanation of a gene-phenotype connection, in molecular genetics practitioners usually consider successful those accounts that contribute enough mechanistic and statistical evidence to establish the link. As we are going to discuss in more detail now, gene position has played a crucial role in both classical and molecular explanations of this sort.

5. Molecular markers

So far, we have shown how the analysis of gene positions exemplifies the molecular retooling of classical genetics. Gene positions play a crucial role in the explanation of mutation effects: the identification of a classical region in the chromosome is the starting point for the molecular analysis of the candidate mutation triggering the mechanisms. To the extent that classical position plays a role in this explanatory pattern, the retooling is, as Waters contends, a practical hybridization of classical and molecular approaches. Our next step is to show how a molecular technique makes classical gene position dispensable. For this we need to introduce the concept of the *molecular marker*.

(Paterson et al. 1988) used molecular markers to identify phenotype-associated genomic regions in tomato lines. Markers, as we are going to see next, allow geneticists to identify molecular positions in a way that makes the classical gene concept dispensable in the analysis. Markers emerged in the study of linkage in classical genetics. Linkage (Darden 2006; Lobo and Shaw 2008) is an exception to the so called second law of Mendel, the law of independent assortment. When two genes, affecting two different polymorphic traits, find themselves near each other in the chromosome, we will find some unexpected associations between their observed phenotypes. Classical geneticists used experimental crosses in *Drosophila melanogaster* to observe and count phenotypes of the second generation individuals. Therefore, they used phenotypes as *markers*, as the unexpected associations between phenotypes indicated that the causal genes were located nearby in the chromosome. In this sense, when Morgan set up his experiment on white-eyed *Drosophila*, he used the phenotypic sex of flies as a *phenotypic marker* for eye color. Sex was already associated with the X chromosome, and Morgan conjectured that the locus for white eyes should be located there.

12

Sex, the X chromosome and eye color followed a concordant segregation (or a co segregation) within his experimental crosses.

A *molecular marker* is defined as a genomic region, generally with a known molecular position, displaying polymorphism (Kinghorn and van der Werf 2000) The statistical analysis of this variability leads to the identification of alleles linked to phenotypic diversity, through association studies. The development of the molecular approach has allowed geneticists to identify many new polymorphisms in the genome (from single base changes to big insertions and deletions) and refine their understanding of molecular gene positions. For instance, the invention of the polymerase chain reaction brought about such popular sources of variability in genetics today as microsatellites and SNPs (single nucleotide polymorphisms).

Molecular markers have known positions in the genome. They beacon huge stretches of DNA (rather than single nucleotides), containing hundreds (or even thousands) of genomic elements. Identifying the specific mutation causing the phenotypic effect under study is a cumbersome, but not unfeasible task. The positions identified through molecular markers can then be studied via genome databases to list all the such elements located nearby. These are going to be *positional candidates* for the mechanism of phenotypic expression. The list can be shortened focusing on the so- called *functional candidates*, which may have a role in significant metabolic pathways and that should then be assayed in molecular experiments. Ultimately, a single candidate mechanism for the mutation should emerge. Its position is identified with a nucleotide number.

Molecular markers may exhibit variability at different levels, e.g., chromosomic or cytogenetic polymorphisms. Yet, key for their success as a laboratory tool is that most markers are polymorphic sites in the *sequence* of the genome: the exhibited variability involves a very low number of nucleotides. This integrates them well within the molecular genetics framework, centered on the DNA molecule. Establishing gene position as a nucleotide number on the DNA sequence became, already in 1988, a guiding principle for the development of the technique. When full genomes were not available, molecular positions were measured as the distance from the targeted position to one or more molecular markers. With the full genome in sight, molecular positions became nucleotide numbers and the markers provided the tool to identify them. In the end, molecular positions are defined through the DNA sequence, in contrast with classical positions, which were just chromosomic locations.

Thus, with molecular markers the genetic approach shifts to the DNA sequence, in a way that goes beyond Water's characterization. Whereas for Waters the first step in the analysis was to find phenotypic variability in organisms, with molecular markers such variations may also be identified directly on the DNA. Steps 1 and 2 in Water's characterization are here merged in a way that suggests, as we are going to defend next, that the classical dimension of the genetic approach is vanishing. As we argued above, an accomplished molecular analysis yields a functional description of the mutation mechanisms involved plus the relevant biochemical analyses at all the relevant molecular levels. Thanks to molecular markers this approach loses its classical levers. Molecular markers are fundamental to gather statistical evidences in the molecular context. The genetic approach can be accomplished without any cytogenetic technique and without focusing on any chromosomic feature of the polymorphism under investigation. As we are going to argue next, this constitutes a form of reduction.

6. Practical reduction

Genetic reductionism is not a popular position in philosophy. It is almost consensual that an epistemic (Nagelian) reduction of classical to molecular genetics, as defended in (Schaffner 1969) is impossible. The terms of both theories are often not connectable (Hull 1974; Rosenberg 1985). And molecular genetics will not provide better explanations than classical genetics about intergenerational transmission (Kitcher 1984), questioning the interest of a reductionist deduction.

A complete ontological reduction from classical to molecular genetics seems equally unlikely. Ontological reductions can operate from type to type (classes of entities) or between tokens (particular entities) (Kaiser 2015). All geneticists will admit that classical genes will somehow correspond to DNA sequences, so some form of ontological reductionism should be possible. Some classical gene properties seem type-reducible in molecular terms: mutation and recombination can be described and measured as processes affecting nucleotides on the DNA molecule. As to the former, classical genetics took the gene as the unit of mutation. Molecular genetics has type-*reduced* mutation to changes in the DNA sequence –any kind of change, as small as one nucleotide. As to the latter, classical genetics detected recombination between genes, which would become the unit of recombination. Molecular studies have shown how recombination is related to DNA break and repair and, again, how it can be measured in nucleotides: the smallest recombination unit is small as two but probably expanding to twelve nucleotides (Benzer 1955; Chakraborty et al. 2018).

Yet, even if mutation and recombination are type-reducible, position and function do not admit even a token-reduction. Token classical positions are particular chromosomic locations, whereas molecular positions are generally a statistical aggregate, with many different instantiations. As we are going to see next, molecular positions, estimated through genomic sequencing, yield a nucleotide number. And there is no token-token correspondence between a cytogenetic band on a chromosome and these nucleotide numbers. Although there were nucleotides in the chromosomic bands defining classical positions, molecular positions are statistical types which are usually not instantiated in a predictable manner on a given individual.

In current linkage analyses, molecular positions are calculated through recombination frequencies between molecular markers. These recombination frequencies are an estimate of the actual distance (in nucleotides) between the markers. Recombination frequencies can vary in different parts of the genome (Gabriel et al. 2002), between species and even between sexes (females tend to recombine less than males), thus producing quite different estimates of distance and gene location. Hence, molecular positions obtained through linkage do not have a direct token correspondence with classical positions. It may be objected that linkage analysis originated as a tool in classical genetics, pivoting on phenotypic markers and usually implemented in combination with some 'anchor' cytogenetic positions. Yet the molecular approach transformed linkage taking it to a completely different level of precision. Linkage analyses did not yield complete genomic maps until the development of molecular markers.

In genomic sequencing, molecular positions are obtained as the product of an alignment, an arrangement of several overlapped sequences. These sequences do not need to belong to the same individual (several organisms of the same species may be used to construct its genomic map) and they don't start at the same point. Thus, the number indicating nucleotide position that can be obtained by an alignment is completely dependent on (and relative to) the sequences included. To overcome the difficulty of having different position in each experiment, reference genomes are built. They are a type, a statistical construct that researchers use to consensually identify and label the positions of the genes they are studying.

Therefore, there is no token-token correspondence between the classical position of a gene for a particular trait to a molecular position[6]. This is all grist for Water's mill. Theoretical

---

[6] As Davide Vecchi (personal communication) observes, a token molecular position might be characterisable – independently of whether it iss epistemically accessible – as the gene location in the genome of a specific cell. But, as of today, there is no way to reduce token classical position to the latter

reduction is not a viable option. Layer-cake antireductionism is not completely defensible either: some classical gene properties are ontologically reducible to molecular terms. If we focus instead on actual genetic practice we see instead how classical approaches are molecularly retooled, with concepts from both approaches coexisting in the implementation of the genetic approach. For Waters, the reduction in a reductive retooling is just a recasting without any actual implication for the underlying theories or models. As we have shown in the previous two sections, such retoolings actually happen in genetics, but they may be just an episode in a bigger reductive process. Waters characterizes the genetics approach as a method. If we focus instead on the outcome, as we suggested in the previous two sections, we find, in addition, explanatory patterns that allow geneticists to decide about the investigative steps to implement in the genetic approach. To the extent that an element becomes dispensable, reduction becomes something more than a simple recasting.

Given an epistemic benchmark, such as the explanatory pattern for phenotypic traits discussed in sections 4-5, we will define practical reduction as follows. Let's assume a technique X, articulated with concepts from an approach A, allows scientists to accomplish the epistemic benchmark α. If this same α is achieved by a technique Y, articulated with concepts from an approach B, and scientists prefer Y to X, X will have been practically reduced. Behind technical change there are indeed many potential drivers, so we will remain here neutral as to the epistemic costs and benefits of such a replacement. Our only goal is to highlight how concepts disappear with the replacement of the techniques they contribute to articulate.

Let us compare our concept of practical reduction with two other similar forms of reduction discussed so far in the literature: *successional reductions* and *empirical equivalences*. (Nickles 1973) and (Wimsatt 1976) introduced the concept of *successional reduction* to account for the replacement of less mature for more mature theories about the same domain of phenomena. Successional reductions hinge on some sort of conceptual correspondence between the theories involved: any theoretical or observational statement of the precedent theory can be expressed in the posterior theory. Following Waters' lead, our concept of practical reduction focuses instead on the research process and how a sequence of different methods may achieve the same epistemic benchmark. Unlike in successional reductions, we do not expect any particular correspondence between previous and posterior theories. Against Waters, we

---

entity. On the other hand, according to (Weber 2004), the situation is equally complicated for reducing function, the fourth characteristic of the classical gene concept, although we'll skip the discussion here.

expect instead that those methods associated with the more mature theories will replace previous methods associated with the antecedent theories. This sets a limit to the pluralism of scientific practices: theories would be guiding the choice of methods more than Waters seems to admit.

A practical reduction is also different from what the structuralist tradition calls an *empirical equivalence* (Balzer et al. 1987). Empirically equivalent theories are those for which there is no correspondence between their concepts or laws, and yet, e.g., the same predictions obtain in both. For instance, Lagrangian and classical mechanics (Balzer, 1987: 292-295}. The equivalence between theories is empirical inasmuch as it connects their data models, the empirical structures (without theoretical terms) both theories target. For Waters, a reductive retooling occurs when an investigative approach incorporates new methods guided by practical, and not theoretical considerations. Thus, we may expect that old and new methods when applied to the same problem will yield empirically equivalent models: whatever the correspondence between their concepts, they deliver similar enough results.

However, the structuralist reconstruction of classical genetics shows that the concept of gene position is theoretical, not empirical. As we already saw, it is part of the definition of the classical gene (Weber 2004), and it features accordingly in some of the key conceptual models of classical genetics (Balzer and Lorenzano 2000). Therefore, the transition between classical and molecular methods for dealing with gene positions will generate something more than an empirical equivalence, it will have theoretical implications that the concept of retooling fails to grasp. The concept of practical reduction highlights how some theories become dispensable when the methods they support become replaceable. Even if a theoretical reduction between classical and molecular genetics is still unattainable, there is a reductive (eliminative) dimension to the sequence of scientific practices. Practical reductions are not about methods disappearing, but rather about the dispensability of some old theories, once the methods they support are replaced.

For position, molecular markers illustrate this shift. Classical positions were identified mainly with two techniques: cytogenetics and linkage analyses with phenotypic markers. With the former, genes were found in chromosomic bands, whereas with the latter genes were identified by proximity. As we have seen, these two techniques are conceptually articulated on the basis of classical genetics: they locate *genes* in the genome, as part of a research process in which geneticists try to gather any piece of information that may be potentially relevant to

explain the observed phenotypic effects[7]. Thus, having both chromosomic information and gene order was helpful to describe patterns of inheritance.

Both cytogenetics and linkage analysis were crucial for implementing Waters' genetic approach by identifying sources of phenotypic variation. Both techniques have been now updated, especially with the use of molecular markers: putative causal mutations are identified through neighbouring polymorphisms that inform about their (approximate) location on the chromosome. As we have already seen, molecular genetics protocols start with the identification of these candidate mutations that are subsequently assessed with other functional studies. These approximations based on the use of molecular markers do not depend on the classical conceptual framework, they are entirely articulated in molecular terms. Position provides the site in the genome on the basis of which geneticists will define the sequence and the polymorphisms that are going to be further investigated as causes of the phenotype.

Therefore, classical gene positions have become practically dispensable once molecular markers allowed geneticists to identify positions without presupposing any classical concept. A quick rejoinder dismissing the relevance of this sort of practical reduction would appeal to its lack of theoretical significance. Some techniques have been replaced, yes, but the epistemic and ontological reduction of classical gene position has nonetheless been achieved. We disagree. In our view, the practical dispensability of classical gene position shows something more: it played no fundamental theoretical role in classical genetics. Gene position was a feature of the classical gene to the extent that it was crucial for the practical implementation of the genetic approach. But it had no causal role in articulating the concept of the classical gene.

The strongest conceptual contribution of gene position to classical genetics could have been causal: gene positions would have an influence on certain phenotypic effects. This idea was indeed explored and discarded, especially when studying the *Bar eye* in *Drosophila*. The *Bar*

---

[7] We are, of course, simplifying our discussion of classical genetics. Following Ken Waters' suggestion (personal communication), it may be argued that recombination maps may exhibit more continuity between classical and molecular approaches than our analysis claims. Recombination maps based on phenotypic markers were used to establish recombination distances between *loci*, and these distances were, in turn, related to distances between chromosomic positions. Molecular geneticists can now build recombination maps using genetic markers: with different data, they can calculate distances with statistical approaches similar enough to those of classical genetics. However, in our view, the relevant difference lies in the source data: unlike in classical genetics recombination maps are now build upon molecular information. The positions under discussion are also different, just as our general argument suggests.

eye phenotype is an alteration of the regular eye which was firstly observed in 1913 by S. C. Tice (Carlson 1966). It is caused by duplications in a region of the X chromosome, and several shapes have been described and related to the numbers of duplications and their chromosomic location (in females, in which of the two homologous chromosomes they are found). After several studies, it was evident that *Bar* displayed quite interesting peculiarities and it was further studied in Morgan's laboratory. In 1925, Sturtevant discovered that the relative position of identical genes affected their action, and he suggested the hypothesis of position effect (Sturtevant 1925). He was the first defending the view that the phenotype might be altered by changes in the position or chromosomal environment of the gene. A few years later, (Thompson 1931) suggested a model for such location with his Episome-Protosome model, suggesting that these two different entities are part of the gene. Then, (Muller 1932) was able to cytogenetically demonstrate that *Bar* was associated to duplications of a chromosomic region caused by an unequal crossing over, and defended again the hypothesis of the position effect (Muller 1936), as the effect of *Bar* would depend not on the number of copies of the chromosomic region, but in their organization in the two homologous chromosomes. From 1937 to 1954, Richard Goldsmicht used *Bar* and position effects to attack the gene concept that was being constructed (Goldschmidt 1937). He defended the importance of the chromosome as the unit controlling normal development, while the changes in the correct order would produce the *deviant* mutants. He thought these order alterations were the so far described position effects, and criticized point mutations and the idea of the gene itself.

Eventually, it became clear that these positions effects were not widespread in the genome. X-ray exposure was shown to mimic the effect of spontaneous mutations (Muller and Muller 1930), which were visible as chromosomic rearrangements, and radiation experiments were performed to assess the effects of new mutations. Point mutations were demonstrated to account for a high number of mutant phenotypes and not even classical geneticists kept arguing for position being fundamental in phenotypic variation. Mutation took the lead, and with the discovery of the structure of the DNA molecule and the articulation of the concept of molecular gene, nucleotide sequence gained all the relevance. In this sense, transgenesis gives further proof that position is not a fundamental characteristic of the gene: transgenes may work even if they are inserted randomly in the genome. Interestingly, it must be acknowledged that,

although position effects for several phenotypes do exist (including *Bar*), currently the underlying mechanisms are usually suspected to be epigenetic.[8]

7. Concluding remarks

Reductionism about theories seems to go against pluralism: if scientists are pursuing successful theoretical reductions, the number of competing theories within a field or in between neighbour fields would decrease, sooner or later. One account would emerge encompassing them all. Scientific practices seem to provide instead a fruitful domain for epistemic pluralists. As (Waters 2019) puts it, in a world without a single fundamental structure, in which scientists work with multiple goals in sight, scientific knowledge is bound to be pluralistic. That is, again in (Waters 2019) own words, "partial, piecemeal, and perspectival". In the actual practice of science different conceptual frameworks may coexist in ways that would be puzzling for the theoretical reductionist. For a reductionist, the coexistence of classical and molecular concepts of gene position over several decades in the laboratory is, surely, an anomaly that the advancement of genetics should correct. For a pluralist, instead, there is no a priori reason to expect such a correction: scientific traditions can be abandoned or taken up for a number of *a posteriori* reasons, without any implication about the structure of the world (Chang 2007)

However, we may wonder whether the introduction of molecular markers has made our understanding of gene positions more or less "partial, piecemeal, and perspectival". Like (Chang 2007), our account of the replacement of classical gene positions for molecular markers is about epistemic progress within a particular tradition. We have argued that when molecular geneticists study the connection between gene regions and phenotypic traits, they consider a connection successfully explained when they gather enough statistical and mechanistic evidence to support it. This benchmark arises from the very practice of genetics, rather than from any external philosophical desideratum. Moreover, it has been widely shared by both classical and molecular geneticists. Here is why we think that there is an internally reductive dimension in the introduction of molecular markers: all the necessary evidence for explaining the connection between genes and phenotypes could be gathered on purely molecular grounds, without any

---

[8] Ken Waters (personal communication) objects that we are overlooking the role position played in the recombination of traits through transmission, a central role for the experimental practices of classical geneticists. Indeed, classical genetics were more successful in the study of recombination than in phenotype explanation (as compared to molecular genetics). Although crucial, position in the study of recombination is *relative* (as in "how separated are these two genes on the genome?"). It does not have the sort of causal impact attributed to position effects.
.

classical tools. In our view, classical geneticists would have appreciated the advantages of such practical advances, as they can be used to gather information about fundamental aspects of the classical gene. Moreover, they are extremely useful to address the mechanisms underlying phenotypic expression, starting with the detailed characterization of the involved genomic region. The description of such variation has always been a main topic of genetics since its foundation.

The dispensability of classical gene position brings about no epistemic or ontic reduction: there is still no token-token correspondence between classical and molecular positions and there is, of course, no straightforward theoretical reduction between them. However, the practical reducibility of gene positions has some consequences for our understanding of the fundamental structure of the world: classical gene positions were not part of it, they had no causal role in the production of phenotypes. Classical positions were just instrumental traits for gene manipulation, completely unlike traits like recombination and mutation.

As (Waters 2019) shows, from the perspective of scientific practices, genetics is in many ways more "partial, piecemeal, and perspectival" than it was, say, a century ago. But our case study illustrates instead that the practice of genetics also brings about conceptual unification, making classical concepts dispensable and showing that they had less ontological traction than geneticists originally expected –probably, none. In this regard, there has been more than a practical retooling.

Of the two characteristics of the classical gene, two have been explained in molecular terms (recombination and mutation), and position has been made practically dispensable. It remains to be seen what the fate of classical gene functions will be. Suppose that the latter became practically reducible in a way that we cannot anticipate today: the advancement of genetic practices would have then achieved a sort of conceptual unification of the field in which all the causal levers would be molecular. This is a wild guess, of course. Its point is to suggest that focusing on scientific practices may not necessarily highlight the internal disunity of science, it can also bring about reductions through other means.

REFERENCES

Amberger JS, Bocchini CA, Scott AF, Hamosh A (2018) OMIM. org: leveraging knowledge across phenotype–gene relationships Nucleic acids research 47:D1038-D1043

Andersson L et al. (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs Science 263:1771-1774

Baetu TM (2010) The Referential Convergence of Gene Concepts Based on Classical and Molecular Analyses International Studies in the Philosophy of Science 24:411-427

Balzer W, Lorenzano P (2000) The Logical Structure of Classical Genetics Journal for General Philosophy of Science 31:243-266

Balzer W, Moulines CU, Sneed JD (1987) An architectonic for science : the structuralist program. D. Reidel, Boston, MA.

Benzer S (1955) Fine structure of a genetic region in bacteriophage Proceedings of the National Academy of Sciences of the United States of America 41:344

Bhattacharyya MK, Smith AM, Ellis TN, Hedley C, Martin C (1990) The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme Cell 60:115-122

Bridges CB (1914) Direct proof through non-disjunction that the sex-linked genes of Drosophila are borne by the X-chromosome Science 40:107-109

Carlson EA (1966) The gene; a critical history. Saunders, Philadelphia,

Chakraborty P, Pankajam AV, Dutta A, Nishant KT (2018) Genome wide analysis of meiotic recombination in yeast: For a few SNPs more IUBMB Life 70:743-752

Chang H (2007) Scientific Progress: Beyond Foundationalism and Coherentism 1 Royal Institute of Philosophy Supplements 61:1-20

Darden L (2006) Reasoning in biological discoveries : essays on mechanisms, interfield relations, and anomaly resolution. Cambridge University Press, Cambridge ; New York

Falconer D (1962) Sombre (So) on LG XVIII Mouse News Lett 27:30

Gabriel SB et al. (2002) The structure of haplotype blocks in the human genome Science 296:2225-2229

Gehring WJ, Klemenz R, Weber U, Kloter U (1984) Functional analysis of the white+ gene of Drosophila by P‐factor‐mediated transformation The EMBO journal 3:2077-2085

Goldschmidt R (1937) Spontaneous chromatin rearrangements in Drosophila Nature 140:767

Green M (1959) Spatial and functional properties of pseudo-alleles at the white locus in Drosophila melanogaster Heredity 13:303

Griffiths P, Stotz K (2013) Genetics and philosophy : an introduction. Cambridge introductions to philosophy and biology. Cambridge University Press, Cambridge, UK ; New York

Hoekstra H (2014) The secret of a natural blond Nature genetics 46:660

Hull DL (1974) Philosophy of biological science. Prentice-Hall foundations of philosophy series. Prentice-Hall, Englewood Cliffs, N.J.,

Kaiser MI (2015) Reductive explanation in the biological sciences. Springer,

Kinghorn B, van der Werf J (2000) Identifying and incorporating genetic markers and major genes in animal breeding programs Curso desarrollado en Belo Horizonte 31

Kitcher P (1984) 1953 and all that. A tale of two sciences The Philosophical Review 93:335-373

Lenffer J et al. (2006) OMIA (Online Mendelian Inheritance in Animals): an enhanced platform and integration into the Entrez search interface at NCBI Nucleic acids research 34:D599-D601

Levis R, O'Hare K, Rubin GM (1984) Effects of transposable element insertions on RNA encoded by the white gene of Drosophila Cell 38:471-481

Lobo I, Shaw K (2008) Discovery and types of genetic linkage Nature education 1:139

Morgan TH (1910) Sex limited inheritance in Drosophila Science 32:120-122

Muller HJ (1932) Further studies on the nature and causes of gene mutations Proc Sixth Int Cong Genet, Ithaca, New York, USA 1:213-255

Muller HJ (1936) Bar duplication Science 83:528-530

Muller HJ, Muller H (1930) Types of visible variations induced by X-rays in Drosophila Journal of Genetics 22

Nickles T (1973) Two concepts of intertheoretic reduction The Journal of Philosophy 70:181-201

O'Hare K, Murphy C, Levis R, Rubin GM (1984) DNA sequence of the white locus of Drosophila melanogaster Journal of molecular biology 180:437-455

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms Nature 335:721

Pierce BA (2013) Genetics essentials : concepts and connections. Second edition. W.H. Freeman and Company, New York

Robbins LS et al. (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function Cell 72:827-834

Rosenberg A (1985) The structure of biological science. Cambridge University Press, Cambridge; New York

Russo F, Williamson J (2007) Interpreting Causality in the Health Sciences International Studies in the Philosophy of Science 21:157-170

Schaffner KF (1969) The Watson-Crick model and reductionism The British Journal for the Philosophy of Science 20:325-348

Searle A (1968) An extension series in the mouse Journal of Heredity 59:341-342

Sturtevant AH (1925) The effects of unequal crossing over at the bar locus in Drosophila Genetics 10:117

Thompson DH (1931) The side-chain theory of the structure of the gene Genetics 16:267

Waters CK (1994) Genes Made Molecular Philosophy of Science 61:163-185

Waters CK (2008) Beyond Theoretical Reduction and Layer‐Cake Antireduction: How DNA Retooled Genetics and Transformed Biological Practice. In: Ruse M (ed) The Oxford Handbook of Philosophy of Biology. Oxford University Press, New York, pp 238–262

Waters CK (2014) Shifting attention from theory to practice in philosophy of biology. In: New directions in the philosophy of science. Springer, Dordrecht, pp 121-139

Waters CK (2019) Presidential Address, PSA 2016: An Epistemology of Scientific Practice Philosophy of Science 86:585-611

Weber M (2004) Philosophy of experimental biology. Cambridge University Press, Cambridge ; New York

Wimsatt WC (1976) Reductive explanation: A functional account. In: PSA 1974. Springer, Dordrecht, pp 671-710