Journal of Law and the Biosciences, 85–91
doi:10.1093/jlb/lsu035
Peer Commentary
Advance Access Publication 23 December 2014

Military genomic testing: proportionality, expected benefits, and the connection between genotypes and phenotypes

Charles H. Pence

Department of Philosophy and Religious Studies, Louisiana State University, Baton Rouge, LA 70803, USA
Corresponding author. E-mail: charles@charlespence.net

ABSTRACT

Mehlman and Li offer a framework for approaching the bioethical issues raised by the military use of genomics that is compellingly grounded in both the contemporary civilian and military ethics of medical research, arguing that military commanders must be bound by the two principles of paternalism and proportionality. I agree fully. But I argue here that this is a much higher bar than we may fully realize. Just as the principle of proportionality relies upon a thorough assessment of harms caused and military advantage gained, the use of genomic research, on Mehlman and Li's view, will require an accurate understanding of the connection between genotypes and phenotypes – accurate enough to ameliorate the risk undertaken by our armed forces in being subject to such research. Recent conceptual work in evolutionary theory and the philosophy of biology, however, renders it doubtful that such knowledge is forthcoming. The complexity of the relationship between genotypic factors and realized traits (the so-called 'G→P map') makes the estimation of potential military advantage, as well as potential harm to our troops, incredibly challenging. Such fundamental conceptual challenges call into question our ability to ever satisfactorily satisfy the demands of a sufficiently rigorous ethical standard.

KEYWORDS: genomic testing, G→P map, genomic regulatory networks, development, proportionality

INTRODUCTION

As with nearly every technological advance, the development of genomics holds out the promise of fruitful military applications. Mehlman and Li persuasively—and, I believe, correctly—argue that the appropriate bioethical framework for the use of genomic science by the military should be derived primarily from two overriding principles.

First comes the traditional (though often disputed) bioethical principle of *paternalism*. While paternalism in the usual medical context may well be a problematic notion,[1] it is clear in the military context that commanders bear a responsibility for the health and safety of those under their command.

Second is an extension of the well-worn military ethical principle of *proportionality*. Proportionality[2] normally applies only to engagements in which civilian casualties can be expected, and states that a military action is permissible only if the harm to civilians and civilian objects is proportional to the military advantage which directly results from that action. Expanding the principle to our own troops, we say that a soldier can be compelled to undergo a risky procedure only if the harm which can be expected to befall her is proportional to the military advantage which can be expected to result.

Proportionality, in Mehlman and Li's extended sense, thus requires that, for every proposed genomic intervention, we estimate both the harms to friendly forces and the military advantage gained against our enemies. Without the ability to gauge these two quantities, it will be impossible to evaluate whether or not an action is indeed proportional. This is not a new problem in military ethics. The difficulty of estimating proportionality has, for example, been raised by Schmitt and Thurnher, leading them to doubt (due to the context sensitivity and subjectivity of 'military advantage') whether proportionality calculations could ever be accurately executed by an autonomous robotic weapons system.[3]

In this commentary, I do not wish to dispute Mehlman and Li's ethical framework. In fact, I think it is quite likely that—situated as it is in the intersection between military ethics and traditional medical ethics, and illuminated by many studies of prior successes and failures of medical research by the military—these two criteria are exactly those which ought to be used to guide discussion of the ethical use of genomics by the military. Rather, I wish to explore in detail one facet of their framework: the estimation of harms and benefits in the context of genomics. Unlike several of the case studies that they discuss, I argue that there are significant *in-principle* (not just research-based or in-practice) problems with the connection between genotypes and phenotypes. These issues will, at the very least, make incredibly difficult any proper assessment of the proportionality of genomic intervention by the military. At worst—and I will argue that there are good reasons to think that we are indeed at worst—these considerations will make it *impossible* to estimate proportionality.

---

1 Robert Young, *Informed Consent and Patient Autonomy*, in A Companion to Bioethics 530, 540 (Helga Kuhse & Peter Singer eds., 2nd ed. 2009).

2 Itself often misunderstood; see Gary D. Solis, The Law of Armed Conflict: International Humanitarian Law in War 280, 283 (2010).

3 Michael N. Schmitt & Jeffrey S. Thurnher, *"Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict*, 4 Harv. Natl. Sec. J. 231, 253–257 (2013); Jeffrey S. Thurnher, *Examining Autonomous Weapon Systems from a Law of Armed Conflict Perspective*, in New Technologies and the Law of Armed Conflict 213, 228 (Hitoshi Nasu & Robert McLaughlin eds., 2014).

I want to stress that I do not intend this to be an argument *against* Mehlman and Li's framework. On the contrary, as I have said above, I believe that Mehlman and Li have set the bar in precisely the correct place. That bar, however, is significantly higher than it may seem on a first reading. Justifying any complex military use of genomics will be—and *should* be, for precisely the reasons that I will outline here—a difficult, if not impossible, enterprise.

THE LONG AND WINDING ROAD FROM GENOTYPES TO PHENOTYPES

The fundamentals of my critique are drawn from recent work in both biology and the philosophy of biology. But first, a brief introduction. An organism is born with a particular genome—the contents of its DNA or 'genetic code'. We refer to this collection of genes as its *genotype.* The genotype, however, is only one of the resources that produce the traits of the organism that we actually see—its *phenotype.* Of course, it is phenotypic traits with which we are actually concerned, as these are the environmentally expressed characteristics, behaviors, morphologies, and so forth that actually matter to the organism during its life.

The crucial question is this: *What is the relationship between genotypes and phenotypes?* The determination of this relationship is one of the most difficult open problems in contemporary biology. Why? An organism's development, the environment into which it is born and in which it later lives, and interactions between vast networks of genes make it difficult to precisely trace the connection between the simple *possession or absence* of a particular gene and the possession or absence of some corresponding phenotype in a one-to-one way. As many commentators have put it,[4] the traditional metaphor of genes as a 'blueprint' for the construction of the organism must be discarded. This has led to the reconceptualization of this connection as the genotype–phenotype map (or G→P map)—the relation from genotypes to phenotypes, which includes all the complex, non-linear interactions of development, genes, and environment.[5]

For some traits in some systems, this complexity can be worked around. In Mendel's original study of pea plants (which we all learn in our high school biology classes), Mendel's data give the distinct impression that the very simple traits of peas with which he was concerned (green v. yellow peas or smooth v. wrinkled skin) can be predicted accurately by merely determining whether or not the plant at issue possesses a particular gene.[6]

THE COMPLEXITY OF INTERESTING PHENOTYPES

However, to understate the issue, a soldier is not a green, wrinkly pea.[7] A DoD report referenced by Mehlman and Li cites as traits of interest phenomenally complex phenotypes such as susceptibility to 'post-traumatic stress disorder, the ability to tolerate conditions of sleep deprivation, dehydration, or prolonged exposure to heat, cold, or

---

4 Massimo Pigliucci, *Okasha's Evolution and the Levels of Selection: Toward a Broader Conception of Theoretical Biology*, 25 Biol. Philos. 405, 415 (2010); Russell Powell, Guy Kahane & Julian Savulescu, *Evolution, Genetic Engineering, and Human Enhancement*, 25 Phil. Tech. 439, 458 (2012).

5 Pere Alberch, *From Genes to Phenotype: Dynamical Systems and Evolvability*, 84 Genetica 5, 11 (1991).

6 Gregor Mendel, *Experiments in Plant Hybridization*, IV Verhandlungen Naturforschenden Vereines Brünn 3, 47 (1866).

7 Even Mendel's data are not so simple, as was recognized at the time; see W. F. R. Weldon, *Mendel's Laws of Alternative Inheritance in Peas*, 1 Biometrika 228, 254 (1902).

high altitude, or the susceptibility to traumatic bone fracture, prolonged bleeding, or slow wound healing'.[8] Later, Mehlman and Li describe 'genetic variants associated with coolness under fire' (p. 28) as a target for possible military genomic screening. Assuming that coolness under fire is a concept sufficiently well defined to count as a phenotype in the first place, it will clearly be a massively complicated one.

How complicated? For the sake of argument, let's say that we are able to *ignore entirely* the influences of development and environment on 'coolness under fire'. This is obviously implausible, but we will be able to make trouble even without considering these factors. To get an idea of the complexity of the genetic regulatory networks that produce traits, consider the work of Davidson et al.[9] They describe a single network which controls the differentiation of two body layers in the embryo of the sea urchin. The resulting diagram contains more than 40 genes, in a highly interconnected, robust, and self-regulating network. And this is a network for a single step in the development of a highly simplified and well-understood model organism. Inference of gene networks in mammals is even more difficult, confounded by more complex regulatory architecture, higher post-transcriptional modification, and the need to refer not only to gene expression data, but also perturbation (knock-down or over-expression) experiments.[10]

Thus far, we've seen a variety of in-practice concerns with the estimation of the G→P map—but as of yet, nothing of an in-principle sort, nothing that would lead us to believe that it is impossible for sufficient research into human systems to solve the problem. After all, if Davidson can describe the genomic regulatory networks for sea urchin embryo development, is it not merely a difference in degree between these networks and those present in humans? I will argue that it is not—with enough distance, a difference in degree becomes a difference in kind. But first, a return to questions of ethics—Why is trouble in the G→P map a difficulty for us here in the first place?

MAKING ETHICAL DECISIONS

For this, we must return to the concept of proportionality. If we are to accurately estimate *either* the harm that might befall our own troops *or* the military advantage that would result from a particular genomic intervention, we clearly must know what phenotypic effects intervening on a particular gene will have in a particular population. And this problem, then, *just is* the problem of understanding the G→P map. Without the ability to translate from genetic to phenotypic effects, estimations of both halves of the proportionality calculus will be hard to come by.

Mehlman and Li recognize this empirical difficulty—they argue that 'the military must be mindful of relying too heavily on the results of genomic tests that have not been adequately validated' (p. 31), mentioning particularly the popular media's tendency to overstate the significance and potential power of genomic results. They explore, then, in some detail, the question of how the military ought to move forward in cases where there is a compelling reason to deploy a genomic technology which

[8] JASON, The $100 Genome: Implications for the DoD 43 (2010), http://www.fas.org/irp/agency/dod/jason/hundred.pdf (accessed Dec. 10, 2014).

[9] Eric H. Davidson et al., *A Genomic Regulatory Network for Development*, 295 Science 1669, 1678 (2002).

[10] Djordje Djordjevic et al., *How Difficult is Inference of Mammalian Causal Gene Regulatory Networks?*, 9 PLoS ONE e111661 (2014).

hasn't been fully tested, drawing an extensive analogy to off-label use of vaccines against chemical and biological weapons, as deployed fairly extensively in the Gulf War and since (pp. 31–33).

This analogy, however, fails to capture the way in which these problems with the G→P map are disconcerting for ethical evaluations of military genomic technologies. In the vaccine case, we have a straightforward instance of an absence sufficient testing data—an intervention which could readily be shepherded through the traditional clinical trial procedure but has not been, as a result of a lack of time, funding, interest from pharmaceutical companies, or what have you.

I think the kinds of concerns already mentioned make it clear that Mehlman and Li understate the problems in testing and validation for genomic interventions of any real complexity. But for the remainder of the paper, I wish to briefly pursue a more troublesome argument. Given the kinds of *in-principle* difficulties present in the notion of the G→P map, I find it likely that it is in fact *impossible* to obtain enough knowledge concerning the connections between genotype and interesting phenotypes to successfully evaluate proportionality.

IN-PRINCIPLE WORRIES IN THE G→P MAP

A plethora of current theoretical and experimental work has indeed improved our knowledge of the G→P map. As chronicled by Pigliucci, studies from computer science have increased our understanding of modularity and network structure, gene networks are far better understood experimentally than they once were, and RNA folding has served as a fruitful model for more complex G→P relationships.[11] But while this work indicates that progress has been made in simple systems, there is room for skepticism beyond such cases. Pigliucci argues that 'a truly satisfactory empirical understanding of G→P relations in complex organisms may [be] forever beyond our grasp because of practical epistemic limitations'.[12]

What are these practical epistemic limitations? In short, while a minor gap in our empirical understanding of the G→P map may only constitute an in-practice experimental difficulty, a large enough number of these 'minor' gaps rises to the level of inability in principle to make accurate predictions. Consider a study on identical twins described by Powell et al.[13] Identical twins offer a fantastic test case for understanding the G→P map. Two such twins, who have (almost) identical genotypes, serve as a natural experiment that can isolate and thus estimate the contribution to a disease (or other phenotype) arising solely from that disease's genetic basis. After statistical analysis in which the study authors estimated the contribution of the genome to some 20 different diseases, they discovered that 'most sequenced patients would not gain any useful information, since their risk' of disease, *even with* full knowledge of their genome, 'would be similar to that of the general population. In the best-case scenario, the majority of patients might be alerted to one or a few disease risks'.[14] Further, this study was

[11] Pigliucci, *supra* note 4, at 563.
[12] *Id.* at 563.
[13] Powell, Kahane & Savulescu, *supra* note 4, at 455.
[14] *Id.* at 455, 456.

in search of disease phenotypes—and on the basis of our genetic studies of cancer, for example,[15] we have reason to think that many disease phenotypes are *much less* complex than those singled out by Mehlman and Li as of potential military interest.

And it is not difficult to see why it is that, 'for the foreseeable future, [empirical charting of the G→P map] seems feasible only for simple instances of G→P'.[16] Estimation of the G→P map requires input from studies of gene networks and evolution,[17] from development, from physiology and morphology,[18] from ecology, from environmental studies, from behavior, and even, in humans, from sociology and studies of culture.[19] As Boudry and Pigliucci put it, 'once we start talking about complex organisms, particularly those characterized by flexible developmental trajectories, all bets are off'.[20]

It is thus likely that, despite the many advances in our understanding of the G→P map in recent years, the shape of this map for organisms of any significant complexity—not to mention humans—is so intricate as to be, at best, far, far beyond our current capacities for estimation, and at worst (which is the more likely outcome) forever beyond our knowledge.

CONCLUSIONS

Mehlman and Li have done a great service in laying out and carefully arguing for a comprehensive approach to understanding the ethics of the use of genomic science by the military. Indeed, in at least some cases—where the phenotypes at issue are relatively simple[21]—the collection of genomic data by the military and screening of our armed forces has the potential to improve health outcomes for our troops, a goal which we all should support. The obstacles to this simpler kind of program are of an in-practice sort; it would require expensive data collection and analysis, a problem that the military is particularly well poised to solve.

But when it comes to phenotypes that are much more complex than this—phenotypes like 'coolness under fire'—difficulties of an in-principle sort arise. The complexity of the relationship between genotypes and phenotypes makes it unlikely that it is possible for us to obtain enough data to accurately estimate the potential harm to our troops and the potential military advantage that would accrue from a genomic intervention. It will, therefore, be impossible to estimate the proportionality of interventions like these. On Mehlman and Li's framework, then, we cannot determine, for many traits

---

15 Themselves not very successful; see C. Glenn Begley & Lee M. Ellis, *Drug Development: Raise Standards for Preclinical Cancer Research*, 483 NATURE 531, 533 (2012).

16 Maarten Boudry & Massimo Pigliucci, *The Mismeasure of Machine: Synthetic Biology and the Trouble with Engineering Metaphors*, 44 STUD. HIST. PHIL. BIOL. BIOMED. SCI. 660–668, 664 (2013).

17 Stefano Ciliberti, Olivier C. Martin & Andreas Wagner, *Innovation and Robustness in Complex Regulatory Gene Networks*, 104 PROC. NAT'L. ACAD. SCI. 13591, 13596 (2007).

18 Arne B. Gjuvsland et al., *Bridging the Genotype–Phenotype Gap: What Does it Take?*, 591 J. PHYSIOL. 2055, 2066 (2013).

19 Necessary because humans are shaped by their cultural and intellectual environments. For a survey of all these factors, see Kevin N. Laland et al., *More on How and Why: Cause and Effect in Biology Revisited*, 28 BIOL. PHIL. 719, 745 (2013).

20 Boudry and Pigliucci, *supra* note 16 at 664.

21 For example, testing for rare, monogenic diseases; see Powell, Kahane & Savulescu, *supra* note 4, at 455.

(many of the traits in which the military seems to be most interested, at least according to the documents which Mehlman and Li cite), whether or not such an intervention is ethical.

Far from a problem with their analysis, however, I believe that Mehlman and Li have offered a framework that encourages an appropriate level of caution regarding these types of experiments. For complex, human phenotypes, it *should* be difficult to ethically justify such an intervention, and it is a virtue of their ethical framework that it lets us see precisely why.