Commentary

Distributing the future: The weak justifications for keeping human genomic databases secret and the challenges and opportunities in reverse engineering them

Misha Angrist a,⁎, Robert Cook-Deegan b,**

⁎⁎ Corresponding author.
⁎ Correspondence to : North Building Box 90222 Durham, NC 27708-0222.
E-mail addresses: misha.angrist@duke.edu (M. Angrist), bob.cd@duke.edu (R. Cook-Deegan).

1. So sue me

Myriad Genetics, founded in 1991 as a spin-off from the cancer genetics epidemiology unit at the University of Utah and initially funded in part by public money, went on to build a multi-billion-dollar business by discovering and patenting two genes that, when mutated, predispose hereditary breast and ovarian cancer (HBOC) (Williams-Jones, 2002; Allison, 2014). While Myriad’s reputation as a competent test provider was generally exemplary and there was no apparent price premium attributable to the patents, the company’s monopoly on the two genes kept patients from obtaining second opinions or confirmatory testing. Moreover, researchers were prevented from returning results on the two genes to research participants (Carbone et al., 2010; Cook-Deegan et al., 2010). In 2009 the American Civil Liberties Union sponsored litigation against Myriad on behalf of twenty plaintiffs (including HBOC patients), seeking to overturn Myriad’s US service monopoly on genetic testing for HBOC. In 2013 the United States Supreme Court ruled that genomic DNA was a product of nature and therefore not patentable (Association for Molecular Pathology et al., 2013), while engineered DNA molecules were eligible to patent. Almost immediately, a spate of other genetic testing firms announced that they would begin testing for the two genes, BRCA1 and BRCA2, that were once the exclusive province of Myriad (Karow, 2013).

But, as Conley et al. describe in their review of the HBOC genetic testing landscape post-Myriad, whatever the legal precedent the Supreme Court established, in the immediate aftermath of the decision the HBOC marketplace only became messier and more confusing (Conley et al., 2014). In the first of its two commercial strategies for HBOC testing post-SCOTUS, Myriad filed suit against most of its new would-be competitors, some of whom countered with a Gene by Gene acquiesced and settled out of court in February 2014 (Allison, 2014; Conley et al., 2014; Sherkow and Scott, 2014). Others have tried to be proactive before launching their own HBOC tests, seeking declaratory court judgments that would allow them to enter the market without fear of litigation (Conley et al., 2014). In all, thus far eight firms have been sued by Myriad, one settled, and several have countersued; the ongoing cases have been consolidated in the US Federal District Court for Utah, Judge Robert Shelby presiding. In all likelihood, the legal wrangling will outlive the first and broadest of Myriad’s surviving patent claims on BRCA1 and BRCA2, which began to expire in 2015.

Ligation and uncertainty ensure a contentious and turbulent HBOC genetic testing market in the near term. But while Myriad’s patent estate may be vulnerable, the company retains a two-decade head start on its competition and a war chest in excess of $250 million (Gleason et al., 2014b). That is why, at least in part, it seems to us that it is not litigation but rather Myriad’s other major post-SCOTUS commercial strategy – to keep its data as a trade secret in the name of “accuracy” (Tucker, 2014) – that is more important and could set a worrisome precedent for the future of precision medicine, which relies on transparency as to how the work was done and broad access to data in order to replicate initial findings and draw robust conclusions about the use of genomics in clinical care (Angrist and Jamal, 2014).

2. The sagacity of opacity?

Myriad’s nearly two decades of control over the BRCA1 and BRCA2 genes allowed it to amass a large proprietary database of variants in these genes (Cook-Deegan et al., 2013). To its credit, company scientists have classified more than 25,000 mutations in cancer-related genes with respect to their pathogenicity (Gleason et al., 2014b). According to Myriad, its rate of variants of unknown significance (VUS), that is, BRCA variants whose pathogenicity (or lack thereof) cannot be determined with high certainty, was down to 2% in 2013–2014 versus 13%
in 2002 (Egginton et al., 2014; Pruss et al., 2014). This is commendable indeed.

What Myriad has not done for ten years, however, is share those variant classifications with the broader scientific community (Tucker, 2014; Cook-Deegan et al., 2013), which means its VUS rate is unverifiable by anyone outside of the company. Why has Myriad declined to share its data? In recent months it has taken the opportunity to slam public databases time and again for their presumptive inaccuracy, lack of oversight/curation and the liability risks attached to using information contained within them to make clinical decisions (Gleason et al., 2014a; Bowles, 2014; Gleason et al., 2014b; Ray, 2014b). Public databases, according to Myriad’s Chief Medical Officer, are not “sufficiently clinic-ready” and their VUS rate is unacceptably high (Tucker, 2014).

Fine. Let’s concede that: 1) public databases harboring cancer-related variants like ClinVar and the Breast Cancer Information Core (BIC) along with the hundreds of locus-specific databases and handful of other genome-wide variant databases have inaccuracies in them — all databases do, of course, because as the science improves, variants will inevitably be re-classified; and 2) there is arguably a greater incentive for commercial laboratories returning results to patients and subject to liability concerns to ensure that variant classifications are accurate for clinical purposes.

But, as we and others have argued elsewhere, the only way these databases will improve, and the promise of personalized medicine will be realized, is through broad data-sharing, not construction of new silos and fortification of existing ones (Field et al., 2009; Angrist and Jamal, 2014; Quackenbush, 2014). Data quality — and by extension, patient care and liability mitigation — improves when the data have more eyes on them, not fewer. Myriad’s withholding data impoverishes the public databases, while sharing data does not hinder Myriad’s use of either its own data or public data. If other labs do as Myriad has, we will have — forgive us — a myriad of private, competitive databases to the detriment of all. Recently, in light of their discovery that family history of breast and ovarian cancer is an inadequate predictor of familial risk on its own (Gabai-Kapara et al., 2014), HBOC pioneer Mary-Claire King and colleagues called for universal screening of BRCA1/2 in women after age 30 (King et al., 2014). Whatever the merits and financial/logistical challenges of such an undertaking (and we should not underestimate the latter), in the near term how feasible would it be to expand HBOC genetic testing by many orders of magnitude while most of the allelic interpretation data remain inaccessible to anyone outside of Myriad?

3. Legerdemain and the public domain

Myriad’s position is that if public databases are not “clinic-ready” then the company will simply take its ball and go home. And so it has.

But of course this is a non sequitur.

There is nothing preventing Myriad from publishing its own mutation data wherever it wants and curated however it wants to whatever databases and handful of other genome-wide variant databases have inaccuracies in them — all databases do, of course, because as the science improves, variants will inevitably be re-classified; and 2) there is arguably a greater incentive for commercial laboratories returning results to patients and subject to liability concerns to ensure that variant classifications are accurate for clinical purposes.

But, as we and others have argued elsewhere, the only way these databases will improve, and the promise of personalized medicine will be realized, is through broad data-sharing, not construction of new silos and fortification of existing ones (Field et al., 2009; Angrist and Jamal, 2014; Quackenbush, 2014). Data quality — and by extension, patient care and liability mitigation — improves when the data have more eyes on them, not fewer. Myriad’s withholding data impoverishes the public databases, while sharing data does not hinder Myriad’s use of either its own data or public data. If other labs do as Myriad has, we will have — forgive us — a myriad of private, competitive databases to the detriment of all. Recently, in light of their discovery that family history of breast and ovarian cancer is an inadequate predictor of familial risk on its own (Gabai-Kapara et al., 2014), HBOC pioneer Mary-Claire King and colleagues called for universal screening of BRCA1/2 in women after age 30 (King et al., 2014). Whatever the merits and financial/logistical challenges of such an undertaking (and we should not underestimate the latter), in the near term how feasible would it be to expand HBOC genetic testing by many orders of magnitude while most of the allelic interpretation data remain inaccessible to anyone outside of Myriad?

4. Reverse engineer agonistes

A couple of years ago, a consortium of advocates, academics and clinical diagnostic laboratories led by Robert Nussbaum at the University of California, San Francisco, launched the Evidence-based Network for Interpretation of Germline Mutant Alleles (ENIGMA) consortium received an NIH stimulus grant to systematically characterize BRCA mutations of unknown significance using a multitude of biological, computational, and other methods (Spurdle
et al., 2012). Its founders include several Myriad collaborators (and BRCA co-disCOVERers) who chose to work instead in the public domain. And investigators with the Global Alliance for Genomics and Health, a consortium of institutions working in healthcare, research, disease advocacy and bioinformatics, are pooling data from all over the world on variants from current sequencing efforts in HBOC (Hayden, 2014) (http://genomicsandhealth.org/our-work/working-groups/clINICAL-working-group/). These efforts are part of a noncommercial grassroots movement to collect BRCA1/2 variants from clinicians and patients, characterize and interpret them, and deposit them in ClinVar and other open-access databases (Lamberton and Terry, 2014); in other words, they are an effort to re-engineer Myriad's database.

While we see these activities as extremely noble undertakings, it is infinitely more important that HBOC survivors, previvors and advocacy groups see them that way and are willing to contribute their data (Tucker, 2014; Lamberton and Terry, 2014). But will they work? In 2013 Nussbaum estimated that he had reconstructed 1.5% of Myriad's BRCA1/2 database (Kolata, 2013). Based on a series of informal conversations, Conley et al. (2014) guessed that, as of mid-2014, the public databases had collectively amassed some 20–25% of Myriad's data. Meanwhile, as of late August 2014, just seven of the 78 labs offering BRCA testing (as listed in both genetests.org and the Genetic Testing Registry [http://www.ncbi.nlm.nih.gov/gtr/]) had committed to depositing their BRCA1/2 data into ClinVar (http://www.free-the-data.org/; last accessed on 20 August 2014).

Why such slow uptake? It is difficult to know for certain, but a couple of factors are likely at work. One is the time and effort involved. Clinical laboratories are busy places and pulling clinical reports for two specific genes among thousands can be a headache. It is noteworthy that Nussbaum's yield increased substantially when he started offering micropayments to labs for sharing BRCA variants (Kolata, 2013).

The other factor, we suspect, is politics. Myriad collaborates with dozens of clinicians and research scientists. While the company might no longer be able to wield genomic DNA patents as a cudgel, it remains the alpha dog in HBOC testing: its own estimate as of August 2014 was that it retained 92% of the HBOC genetic testing market (Gleason et al., 2014b). A number of oncologists and cancer geneticists we have spoken with informally, while sympathetic with the goals of Free the Data, are nevertheless loath to antagonize Myriad. Moreover, they obviously want the best for their patients, and many are content to send samples to Myriad for analysis.

One of the ironies is that the Supreme Court decision means the Myriad service monopoly that enabled the proprietary database to form rested on patent claims granted in error by the US Patent and Trademark Office. Myriad gained from those mistakes until June 2013, and now the database built on those mistakes further benefits Myriad due to inaction among accreditors and payers.

5. Redistributing the future

What to do, then, to make the BRCA1/2 knowledge-base more widely available? Time will likely prove salutary: expirations on the earliest and broadest patents, which begin next year (Conley et al., 2014), may make would-be BRCA variant depositors feel less inhibited about sharing their test reports. And if, as Nussbaum's experience suggests, retrieval of such reports would benefit from some modest subsidy, one could imagine the Free the Data community and a coalition of willing partners soliciting donations from its supporters to expedite such retrievals.

Another promising approach appeals to patients directly to secure their BRCA results. For what it's worth, this has never been more feasible than it is now: As of February 2014, the US Department of Health and Human Services amended CLIA to allow patients access to their completed test reports at their or their representatives' request. While patients can continue to get access to their test results from their doctors, the 2014 rule provides for patient access directly from the clinical laboratory. The rule is not without hurdles: patients must put their request in writing and may be charged for the costs of photocopying, mailing and/or electronic media such as flash drives. But assuming these requirements are met, in most cases the patient's request must be honored within 30 days (Medicare et al., 2014).

Ultimately, Myriad will find that hoarding is not in its own interests. The recent discovery that female carriers of loss-of-function mutations in the gene PALB2 have a risk of developing breast cancer by age 70 of 35% (and in some cases as high as 58%) is another potent reminder that BRCA1/2 mutations are a necessary but insufficient component of the HBOC story (Antoniou et al., 2014). With costs now $1000 for generating whole-genome sequence raw data (Herper, 2014), the technical and financial barriers to comprehensive genetic assessments of HBOC risk are falling away.

Myriad's own MyRisk test of 25 cancer genes is scheduled to supplant its BRCA-based BRACAnalysis® by mid–2015 (Gleason et al., 2014b), and every cancer clinic is experimenting with cancer panels, if not ordering them regularly. Illumina has been in the genomic sequencing business for longer than Myriad, and it has a database with more than twice as many variants as Myriad's across nearly 2000 genes (E. Ramos, personal communication, 22 August 2014). There will be no service monopoly on multiplex cancer panel testing, exome sequencing, or whole-genome analysis; thus, no one firm can create a database as dominant for the human genome writ large as Myriad's is for BRCA1/2 variants.

So long as incentives remain as they are, the value of Myriad's database will dissipate and eventually it will make no sense for the company to expend resources on maintaining its own database of marginal value, but that depends on the rest of the network developing databases and analytical tools equal to the task. In the meantime, it is understandable, but unfortunate for the system as a whole that Myriad has chosen to build a business model around closely held clinical data.

Myriad has long been willing to wear the black hat (Baldwin and Cook-Deegan, 2013). The quality and speed of its BRCA testing services and resulting support from patients and clinicians have made that hat worth wearing. The company's approach may remain financially viable for a while. But in a whole-genome world where disease risk is mediated by a panoply of genes and patients are increasingly mobilized to demand their data, sooner or later private databases will become an maladaptive strategy.

One black mark that is rarely noted is the impact of Myriad's strategic behavior outside the hotly contested and highly lucrative US BRCA testing market. The effects of the battle for US genetic testing affect the whole world. Myriad reported $10 million a week in revenues for BRCA testing in 2013–14: $517 million out of total revenues of $778 million (Anonymous, 2014). That revenue stream, which has produced $2.8 billion in revenues since 2004, drives Myriad's litigation and database secrecy strategies. Women throughout the world are casualties of the dysfunctional US market, which created incentives for data-hoarding. Thus far public sources have failed to build the infrastructure to interpret genomic variants that would obviate such business strategies, so that while women tested in Malawi and Manhattan face the same choices, the ones in Manhattan are much more likely to be able to overcome barriers to follow-up given their access to a higher baseline of both primary and specialty care than women in the developing world. Unlike the medical literature that is curated for world use and available anywhere – to the immense benefit of patients everywhere – genomic variant data are cloaked in secrecy due to the way some US stakeholders are playing the genetic testing game. If you want your BRCA mutation to be interpreted with the best data available, you have to get both the test and the interpretation performed by a US company in Salt Lake City, whether you live in Moscow, Idaho, or Moscow, Russia. The September 2014 decision by the Federal Court of Australia to uphold Myriad's isolated genomic DNA patents (despite an historical lack of enforcement) suggests that the notion that genetic testing requires monopolies even in mature markets is not quite dead.
Cook-Deegan, R., Conley, J.M., et al., 2013. The next controversy in genetic testing: clinical

Field, D., Sansone, S.A., et al., 2009. Megascience: ‘Omics data sharing. Science 326 (5950), 234–236.

Gaba-Kapara, E., Lahad, A., et al., 2014. Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2. Proc. Natl. Acad. Sci. U.S.A. http://dx.doi.org/10.1073/pnas.1415979111 [Epub ahead of print, pii: 201415979].

Gleason, S., Meldrum, P.D., et al., 2014a. Myriad Genetics’ (MYGN) CEO Peter Meldrum on Q3 2014 Results – Earnings Call Transcript, (6 May).

Gleason, S., Meldrum, P.D., et al., 2014b. Myriad Genetics’ (MYGN) CEO Peter Meldrum on Q4 2014 Results – Earnings Call Transcript, (12 August).

Hayden, E.C., 2014. Cancer-gene data sharing boosted. Nature 510 (7504), 198.

Herper, M., 2014. Flatley's law: how one company is creating medicine's genetic revolution. Forbes (20 August, Retrieved from http://search.proquest.com/docview/1559075537?accountid=10598).

Karow, J., 2013. Competition for Myriad heats up as US testing labs launch BRCA tests, Hereditary Cancer Panels. Clin. Sequencing News (26 June, http://www.genomeweb.com/sequencing/competition-myriad-heats-us-testing-labs-launch-brca-tests-hereditary-cancer-pan).

King, M.C., Levy-Lahad, E., et al., 2014. Population-based screening for BRCA1 and BRCA2: 2014 Lasker Award. JAMA. 312 (11), 1091–1092.

Kolata, G., 2013. DNA project aims to make public a company’s data on cancer genes. N.Y. Times (12 April, http://www.nytimes.com/2013/04/13/health/dna-project-aims-to-make-company-data-public.html?pagewanted=all&_r=0).

Lamberton, K., Terry, S.F., 2014. Free the Data. Genet. Test. Mol. Biomarkers 18 (1), 1–2.

Lee, J.M., Ledermann, J.A., et al., 2014. PARP inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies. Ann. Oncol. 25 (1), 32–40.

MacArthur, D.C., Manolio, T.A., et al., 2014. Guidelines for investigating causality of sequence variants in human disease. Nature 508 (7497), 469–476.

Medicare, C.f, Medicaid Services, H.H.S., et al., 2014. CJIA program and HIPAA privacy rule: patients’ access to test reports. Final rule. Fed. Regist. 79 (25), 7289–7316.

Minear, M.A., Kapustij, C., et al., 2013. Cystic fibrosis patents: a case study of successful licensing. LES Nouv 21–30.

Myriad, 2014. Press release: Myriad Genetics announces participation in innovative collaboration to accelerate hereditary cancer research. Globe Newswire (2 June, http://investor.myriad.com/releasesdetail.cfm?releaseid=-851822).

Pruss, D., Morris, B., et al., 2014. Development and validation of a new algorithm for the reclassification of genetic variants identified in the BRCA1 and BRCA2 genes. Breast Cancer Res. Treat. 147 (1), 119–132.

Quackenbush, J., 2014. Learning to share. Sci. Am. 311 (1), S22.

Ray, T., 2014a. Multiple institutions, labs collaborating on registry to study rare hereditary cancer genes. Genome Web Daily News (2 June, http://www.genomeweb.com/clinical-genomics/multiple-institutions-labs-collaborating-registry-study-rare-hereditary-cancer-g).

Ray, T., 2014b. In tackling the VUS challenge, are public databases the solution or a liability for labs? Pharmacogenomics Report. (12 February, http://www.genomeweb.com/clinical-genomics/tackling-vus-challenge-are-public-databases-solution-or-liability-labs).

Rehm, H.L., Bale, S.J., et al., 2013. ACMG clinical laboratory standards for next-generation sequencing. Genet. Med. 15 (9), 733–747.

Sherkow, J.S., Scott, C., 2014. CASE STUDY: Myriad stands alone. Nat. Biotechnol. 32 (7), 620.

Spurde, A.B., Healey, S., et al., 2012. ENIGMA — evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. Hum. Mutat. 33 (1), 2–7.

Tucker, K.L., 2014. Genetics lab refuses to share data that could save lives. J ew. Daily Forw. (12 August, http://forward.com/articles/203739/genetics-lab-refuses-to-share-data-that-could-save/?p=all).

Williams-Jones, B., 2002. History of a gene patent: tracing the development and application of commercial BRCA testing. Health Law J. 10, 123–146.