

**THIS VERSION DOES NOT CONTAIN PAGE NUMBERS.
PLEASE CONSULT THE PRINT OR ONLINE DATABASE
VERSIONS FOR THE PROPER CITATION INFORMATION.**

ARTICLE

PATENTS AS DATA AGGREGATORS IN PERSONALIZED MEDICINE

DAN L. BURK[†]

INTRODUCTION

The role that patents might play in fostering personalized medicine poses a difficult question. The usual account of the patent mechanism is that it confers exclusive rights on an entrepreneur, allowing her to exclude competitors from the marketplace, and so to charge higher prices for the claimed invention, or for licenses to the claimed invention, than would be possible under fully competitive market conditions.¹ The promise of these higher profits provides an *ex ante* incentive for investments in innovation that may yield patentable inventions.² Absent the patent incentive, innovation might be underfunded, or at least might be underfunded on average, in some fields some of the time.³

The actual success of patents in this regard is at best unclear; it is entirely possible that patents may not provide proper incentives, or the incentives may be more costly than the benefit.⁴ But there is some additional potential for this rationale to break down in the context of personalized medicine.⁵ The usual patent story assumes a robust distribution of consumers purchasing the claimed invention in order to recoup the innovator's investment. Personalized medicine is by definition personalized, using molecular diagnostics to identify optimal treatment regimes tailored to a particular individual or perhaps a small group

[†] Chancellor's Professor of Law, University of California, Irvine.

¹ See Dan L. Burk, *Law and Economics of Intellectual Property: In Search of First Principles*, 8 ANN. REV. L. & SOC. SCI. 397 (2012) (summarizing the economic justification for intellectual property).

² *Id.*

³ *Id.*

⁴ *Id.* at 403.

⁵ See Michael J. Meurer, *Pharmacogenomics, Genetic Tests, and Patent-Based Incentives*, in PERSPECTIVES ON PROPERTIES OF THE HUMAN GENOME PROJECT 399, 400-01 (F. Scott Kieff ed., 2003) (modeling incentives for producing personalized drugs and genetic testing); see also Valerie Gutmann Koch, *Incentivizing the Utilization of Pharmacogenomics in Drug Development*, 15 J. HEALTH CARE L. & POL'Y 263 (2012) (discussing non-patent incentives for personalized medicine).

of individuals with similarly idiosyncratic physiology.⁶ Personalized medicine thus seems likely to have extremely small target markets for any given treatment. To the extent that the invention claimed in a patent is tailored to an individual, or a few individuals, the patent promises relatively little in the way of incentives. The prospect of obtaining rights to exclude competitors from marketing an innovation designed for an individual, or even a few individuals, hardly seems sufficient to attract the investment needed to develop that innovation.⁷ This suggests that any role for patents must surely be different from, and perhaps altogether contrary to, the role patents are believed to play in promoting innovation generally.

The problem of single sale returns on investment is not entirely unique to patent law, of course. A version of the same conundrum exists in copyright, where exclusivity is intended to provide similar incentives toward creative expression, but where certain works of fine art – “original” paintings, sculptures, and the like – exist only in single instantiations.⁸ Since the rewards of exclusivity are not spread out over multiple sales, the price for such works typically builds all the potential return on investment into a single sale of the work. However, unlike the case of personalized therapies, in fine art there always remains the possibility of moving to lower-cost mass production of the fixed work – indeed, copyright exclusivity is probably deployed in such cases to *prevent* extensive reproduction of unique creative works, suggesting that preserving single copy “originals” in fine art constitutes a sales strategy, facilitated by exclusive rights, instead of the inherent necessity of unique innovation in personalized medicine.

This is not to say that a particular patent associated with personalized medicine would always be matched to a single particularized therapy. Some patents may be foundational, covering innovation that would be applicable to a wide range of individualized treatments, or a patent may cover a fundamental device or method that is employed in a variety of personalized diagnoses. Some patents may be broad enough to cover families of molecules or suites of diagnostic tests that can be adapted to individual therapies. However, it seems likely that the familiar pattern of pharmaceutical innovation – where a patent provides the market exclusivity to recoup the costs for research and development of a successful pharmaceutical molecule – will be unavailable in the case of individually developed treatments. And the fine art pricing strategy,

⁶ See Christopher M. Holman, *Patent Eligibility as a Policy Lever to Regulate the Patenting of Personalized Medicine*, in PERSPECTIVES ON PATENTABLE SUBJECT MATTER 114, 115-17 (Michael B. Abramowicz, James E. Daily & F. Scott Keiff eds., 2015) (describing personalized medicine).

⁷ See Koch, *supra* note 5, at 272.

⁸ See Daniel J. Gifford, *Innovation and Creativity in the Fine Arts: The Relevance and Irrelevance of Copyright*, 18 CARDOZO ARTS & ENT. L.J. 569 (2000) (discussing intellectual property application to single and limited copy works).

2015]

PATENTS AS DATA AGGREGATORS

of recouping development costs from a single first sale, would be likely to place personalized treatments, like unique works of art, out of the price range of the average consumer.

All of this might suggest that patents have diminished or negligible potential to play a role in personalized medicine. I will suggest here that, whatever role patents may play in the conventional innovation story, they may play a different role where individualized medicine is concerned. Specifically, in the context of personalized medicine, patent exclusivity may serve less as an incentive to invest in new inventions than it might to serve as an aggregator for certain types of ancillary information that will be critical to personalized diagnosis and treatments. In this essay I look at the effect of patents on the collection and application of such non-patentable data related to genetic variation. My vehicle for examining such effects is the testing service for genetic predisposition to cancer which was the subject of the recent Supreme Court decision in *Association for Molecular Pathology v. Myriad Genetics*.⁹ The decision in *Myriad* might be hailed as freeing genetic testing from the grips of patent monopolists, or denigrated as eviscerating the incentive to invest in future diagnostics.¹⁰ But beyond such ideological posturing there remains a set of effects, both positive and negative, that have gone largely unexplored and undiscussed in the arguments over DNA patenting.

MOLECULAR IDIOSYNCRASIES

I begin with a few words about the nature of personalized medicine. Much of the concept of personalized medicine is based on the phenomenon of microheterogeneity; that is, the incidence of individual or idiosyncratic molecular variations from the norm generally found in human populations. Such variations are found in all classes of physiological molecules, particularly the macromolecular categories of proteins, nucleic acids, and polysaccharides.¹¹ In some senses there is actually no “norm” for the molecules that mediate biological functions; all individuals display to a greater or lesser degree structural variance from the ideal or typical molecules contemplated by biological science.¹² For the sake of simplicity and convenience, science begins with similarities, and then proceeds to contemplate differences, although in the process sometimes differences may be overlooked. These variations may be the cause of disease, or they may dictate

⁹ *Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107 (2013).

¹⁰ WILLIAM HOFFMAN & LEO FURCHT, *THE BIOLOGIST’S IMAGINATION: INNOVATION IN THE BIOSCIENCES* 207 (2014) (discussing controversy over the *Myriad* lawsuit); Christopher M. Holman, Mayo, Myriad, and the Future of Innovation in Molecular Diagnostics and Personalized Medicine, 15 N.C. J.L. & TECH. 639, 640-41 (2014) (same).

¹¹ See generally NAT’L ACAD. OF SCIS., INST. OF MED., *MICROHETEROGENEITY OF BIOLOGICAL MOLECULES: REPORT OF A WORKSHOP 1-2* (1991).

¹² *Id.*

the use of individualized pharmacological treatments that are compatible with the structural variations found in a particular individual.¹³

Microheterogeneity derives from a number of physiological sources, including individual deviations in molecular synthesis and idiosyncrasies in post-synthesis molecular modification.¹⁴ One important aspect of microheterogeneity that appears a likely area of focus for personalized medicine in the near future is genetic polymorphism; that is, idiosyncratic variations in key genetic loci.¹⁵ It is common to talk about genes as if they are uniform¹⁶ when in fact a degree of variation is integral to their function – genetic variation is a key mechanism driving natural selection and species survival. At the same time, some variations in genetic structure or content can be detrimental, leading to malformed, dysfunctional, overproduced, or underproduced gene products, resulting in disease.¹⁷ Different individuals will carry different versions of particular DNA sequences, and identifying which version of a gene or other genetic locus that a patient carries may be critical to diagnosis of a genetically influenced malady, and may in some cases prove to be useful or important in selecting therapies.¹⁸

Developing tests to recognize individualized variations in key genetic sequences, and correlating such variations with particular disease states and treatments necessarily requires assembly of an extensive catalog of idiosyncratic alleles. The patent eligibility of such genetic sequences was the core issue under consideration in the recent Supreme Court decision *Association for Molecular Pathology v. Myriad* where the Court considered whether nucleotide sequences are excluded from patentable subject matter as products of nature.¹⁹ Or, as the Supreme Court formulated the question, are human “genes” patentable subject matter?²⁰ The specific genes at issue in the case were the sequences for BRCA1 and BRCA2, which are genetic loci related to certain forms of breast and ovarian cancer.²¹ Myriad Genetics, which had characterized the consensus BRCA1 and BRCA2 sequences, held patents on both the isolated molecules and the process of using the molecules in a

¹³ *Id.* at 7.

¹⁴ *Id.* at 5-6.

¹⁵ See Joanna L. Mountain, *Personal Genomics*, in I GENOMIC & PERSONALIZED MEDICINE 74 (Geoffrey S. Ginsburg & Huntington F. Willard eds., 2d ed. 2013).

¹⁶ And, for that matter, as if the term “gene” has a settled meaning. See Dan L. Burk, *Edifying Thoughts of a Patent Watcher: The Nature of DNA*, 60 UCLA L. REV. DISC. 92, 95-97 (2013) (discussing the evolving and unsettled application of the term).

¹⁷ See NATIONAL ACADEMY OF SCIENCES, *supra* note 11, at 4.

¹⁸ See Steve Hook, *Molecular Biology*, in A HANDBOOK OF BIOANALYSIS AND DRUG METABOLISM 326, 349 (Gary Evans ed., 2004).

¹⁹ 133 S. Ct. 2107 (2013).

²⁰ *Id.* at 2116.

²¹ *Id.* at 2112.

2015]

PATENTS AS DATA AGGREGATORS

diagnostic test.²² The latter process patents were invalidated by lower courts²³ and only the product patents on the genes were before the Supreme Court.

The translational products of BRCA1 and BRCA2 are proteins that are key contributors to cellular repair systems; they are active in repairing damaged DNA.²⁴ DNA damage occurs with regularity, requiring routine maintenance.²⁵ Failure to repair such damage can lead to abnormal or uncontrolled cellular growth, which is to say cancer.²⁶ When the repair mechanism is itself inoperative or ineffective, unable to fix damaged DNA, the probability of cancer goes up. BRCA1 and BRCA2 are not themselves causes of cancer; rather, mutations that render them inoperative raise the risk of cancer from unrepaired damage.²⁷ Identifying specific mutations in the genes, and linking them with their associated statistical incidence of disease allows patients carrying those particular mutations to know more about their personal risks for developing cancer. Such associations were the basis for a useful clinical test.

The actual mechanism of the test was based on the comparison of patient samples to variations in the BRCA1 and BRCA2 genes, which is to say to particular nucleotide sequences. As I have pointed out elsewhere, the term “gene,” highlighted by the Court’s formulation of the issue, is somewhat ambiguous even in the scientific community that studies such structures.²⁸ In the *Myriad* case itself, there were in fact numerous molecules covered by the patent claims; the molecules at issue fell broadly into two general classes of nucleotide sequences. The first type, genomic DNA or gDNA, constituted DNA molecules as extracted from human cells, essentially the nucleotide sequence as excised from the chromosome.²⁹ A second type of molecule, complementary DNA or cDNA, constituted a nucleotide synthesized under laboratory conditions in a process called reverse transcription, using information from mRNA transcripts.³⁰

²² *Id.* at 2113-14.

²³ *Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 689 F.3d 1303, 1334 (Fed. Cir. 2012) (affirming patent ineligibility of Myriad diagnostic process claims), *aff’d in part, rev’d in part*, 133 S. Ct. 2107 (2013).

²⁴ J.A. Duncan, J.R. Reeves & T.G. Cooke, *BRCA1 and BRCA2 Proteins: Roles in Health and Disease*, 51 *MOLECULAR PATHOLOGY* 237 (1998).

²⁵ Dana Branzei & Marco Foiani, *Regulation of DNA Repair Throughout the Cell Cycle* 9 *NATURE REVIEWS MOLECULAR CELL BIOLOGY* 297 (2008).

²⁶ Suzanne Clancy, *DNA Damage and Repair: Mechanisms for Maintaining DNA Integrity*, 1 *NATURE EDUCATION* 103 (2008).

²⁷ Kiyotsugu Yoshida & Yoshio Miki, *Role of BRCA1 and BRCA2 as Regulators of DNA Repair, Transcription, and Cell Cycle in Response to DNA Damage*, 95 *CANCER SCIENCE* 866 (2004).

²⁸ *See* Burk, *supra* note 16, at 95.

²⁹ JAMES D. WATSON ET AL., *MOLECULAR BIOLOGY OF THE GENE* 39 (6th ed. 2008).

³⁰ *Id.* at 749.

In a dichotomous and profoundly contradictory opinion, the Supreme Court held that Myriad's claims to genomic DNA (gDNA) sequences for the BRCA1 and BRCA2 genes were invalid, lying outside of patentable subject matter, whereas Myriad's claims to complementary DNA (cDNA) sequences of the genes constituted patentable subject matter.³¹ The rationale of the opinion remains a bit of a puzzle, as the genomic sequence patents were invalidated on the ground that they carried the same coding information – the same nucleotide sequence – as the native human chromosome, even if the extracted molecules were structurally altered by virtue of having been extracted.³² At the same time, the complementary DNA molecules were held to constitute patentable subject matter because they differed structurally from the native chromosome, lacking the intervening sequences that would be found in the chromosomal version.³³

The rationales for the respective holdings on genomic and complementary DNA are thus entirely unclear, as are the precise implications for patenting of other biological molecules.³⁴ One might read the opinion as saying patent eligibility is judged by structure, or that it is judged by sequence, or perhaps by some unarticulated combination of the two. What is clear from the opinion, and what the Patent Office has incorporated into its subsequent interim examination guidelines, is that genomic DNA sequences are not patentable subject matter.³⁵ Neither does it appear that process claims involving DNA sequences are patentable subject matter: in a previous case, *Mayo v. Prometheus*, the Court held that many, perhaps all, diagnostic methods are unpatentable, at least to the extent that they rely upon a principle or phenomenon of nature.³⁶ Lower courts have interpreted this to mean that process patents drawn to the diagnostic use of Myriad's genomic DNA sequences are invalid.³⁷

This set of decisions seems to leave Myriad, and firms like Myriad, almost

³¹ Dan L. Burk, *Are Human Genes Patentable?*, 44 IIC – INT'L REV. INTELL. PROP. & COMPETITION L. 747 (2013).

³² Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 133 S. Ct. 2107, 2118 (2013).

³³ *Id.* at 2119.

³⁴ See Burk, *supra* note 31, at 748-49.

³⁵ See Memorandum from Andrew H. Hirshfeld, Deputy Comm'r for Patent Examination Policy, U.S. Patent & Trademark Office, to Patent Examining Corps (June 13, 2013), available at <http://patentdocs.typepad.com/files/uspto-myriad-memorandum.pdf> (archived at <http://perma.cc/4TT9-LJP5>).

³⁶ *Mayo Collaborative Servs. v. Prometheus Labs. Inc.*, 132 S. Ct. 1289 (2012).

³⁷ *In re BRCA1- and BRCA2-based Hereditary Cancer Test Patent Litigation*, 774 F.3d 755, 763-64 (Fed. Cir. 2014); *Ass'n for Molecular Pathology v. Myriad Genetics, Inc.*, 689 F.3d 1303, 1334 (Fed. Cir. 2012), *aff'd in part, rev'd in part*, 133 S. Ct. 2107 (2013)

2015]

PATENTS AS DATA AGGREGATORS

entirely bereft of patent coverage for their core diagnostic technologies.³⁸ In the particular case of the BRCA1 and BRCA2 genetic tests, removal of patent exclusivity quickly resulted in competitors entering the field with lower cost tests.³⁹ The availability of alternate testing sources, and competitive pricing, represents an immediate benefit for patients who need the test. But recall the justification for patenting that I set out earlier: patents temporarily suppress competition in order to promote investment in technologies that might otherwise be under-produced. The lack of both product and process patents in medical diagnostics raises a general concern that without meaningful patent protection, commercial firms may be unwilling to invest in new diagnostic methods, and particularly new genetic diagnostic methods.

AGGREGATING ALLELIC DATA

But the phenomenon of microheterogeneity, on which concepts of personalized medicine are based, may also lead to an alternative strategy for recouping innovation costs. The Myriad Genetics situation appears to be a case in point. The BRCA1 and BRCA2 genes are highly polymorphic, often appearing in different individuals with nucleotide deletions or substitutions at multiple sites; and so the genes consequently exist in a large number of highly variable alleles.⁴⁰ Different mutations lead to the production of proteins that are impaired to a greater or lesser extent; the more impaired the protein, the less DNA repair it can perform, and the greater the risk of unrepaired DNA damage that may lead to cancer. Thus certain mutations, which lead to a highly inoperative protein, carry a higher risk of cancer, and certain mutations, which lead to a more functional protein, carry a lesser risk of cancer.

This is of course the point of developing the diagnostic test: the various alleles have different particularized associations with the risk of cancer. The more alleles a testing laboratory knows about, and the more information it has associating particular variations with particular health outcomes, the more complete and useful the outcome of the test will be. During the course of its patent exclusivity, Myriad tested samples from thousands of individuals, and so accumulated data about the different variations of the two patented BRCA genes.⁴¹ Such a database of variants is of course a valuable resource in

³⁸ See Rebecca S. Eisenberg, *Diagnostics Need Not Apply*, 21 J. SCI. & TECH. L. 256 (2015).

³⁹ See Turna Ray, *Competition Coming for Myriad's BRCA Test, Whether or Not Gene Patents Hold Up*, GENOMEWEB (Apr. 19, 2013), <https://www.genomeweb.com/clinical-genomics/competition-coming-myriads-brca-test-whether-or-not-gene-patents-hold> (archived at <http://perma.cc/4GK4-NFVK>).

⁴⁰ M. William Audeh, *Genetic and Environmental Factors in Cancer Pathogenesis*, in PRINCIPLES AND PRACTICE OF SURGICAL ONCOLOGY 29, 39 (Howard Silberman & Allan W. Silberman eds., 2010).

⁴¹ See Robert Cook-Degan et al., *The Next Controversy in Genetic Testing: Clinical*

identifying the risk associated with different mutations of the gene. But access to the collected information is restricted; despite invalidation of the patent, Myriad still holds its variant data proprietary as a trade secret.⁴² This strategy effectively provides a new, alternative basis for exclusivity over the most comprehensive form of the diagnostic test, to some extent negating the anticipated competition that might result from the Supreme Court's holding regarding the BRCA1 and BRCA2 patents.

Proposals have been made for re-constructing the allelic database as an openly available resource that could be accessed by new testing competitors entering the BRCA1/BRCA2 marketplace.⁴³ Perhaps the most viable proposal for constructing such a parallel resource has been to encourage patients who were previously tested by Myriad to submit the results of their past tests to a single location for inclusion in an open archive.⁴⁴ But any re-construction of the Myriad database is bound to be costly and uncertain; soliciting voluntary proffer of past tests likely means piecemeal development, dependent on the altruism and attention of multitudes of patients who hold individual snippets of data. Additionally, there is no guarantee that going forward such a new database will be updated by the competing testing laboratories that in the absence of patent exclusivity will now process testing samples, raising the possibility that any new voluntary database will rapidly become obsolete.

Indeed, the new competitive environment for the diagnostic test likely means future data on allelic variations will be scattered across multiple small, incomplete databases. Now that Myriad's exclusivity no longer poses an obstacle to market entry, multiple testing firms are already vying to enter the market, competing down the price of the test.⁴⁵ Competition, as we have noted, offers price benefits to patients, or at least to their insurers that are looking for cost savings. But competition also means that data about any given patient's genetic variation may be lodged at any of a number of firms, rather than all being submitted to one monopolist firm. Naturally there could be benefits to

Data as Trade Secrets?, 21 EUR. J. HUM. GENETICS 585 (2013).

⁴² *Id.* at 587.

⁴³ See Misha Angrist & Robert Cook-Degan, *Distributing the future: The weak justifications for keeping human genomic databases secret and the challenges and opportunities in reverse engineering them*, 3 APP & TRANSLATIONAL GENOMICS 124 (2014) (describing multiple efforts to re-construct the Myriad gene variation database); John Conley, Robert Cook-Degan & Gabriel Lázaro-Muñoz, *Myriad After Myriad: The Proprietary Data Dilemma*, 15 N.C.J.L. & TECH. 597, 618-19 (2014) (same).

⁴⁴ See Katherine Lambertson & Sharon F. Terry, *Perspective: Free the Data*, 18 GENETIC TESTING & MOLECULAR BIOMARKERS 1 (2014) (describing the "Free the Data" grassroots genetic testing results collection effort).

⁴⁵ See GenomeWeb, *GeneDx Files IPR Petitions at USPTO against 11 Patents at Issue in Myriad Lawsuit*, GENOMEWEB (Aug. 19, 2014, 2:49 PM), <https://www.genomeweb.com/clinical-genomics/genedx-files-ipr-petitions-uspto-against-11-patents-issue-myriad-lawsuit> (archived at <https://perma.cc/3NZV-C53Q>).

cooperation that might prompt collaboration by the newly competing firms: by agreeing to share data regarding the samples they test, they could gain access to useful data regarding the samples that their competitors have tested. However, while such sharing produces a benefit to any given firm, it also means benefitting rivals who are competing for market share. Firms may have difficulty overcoming the classic “prisoner’s dilemma” incentive to withhold rather than share such data with their direct competitors.⁴⁶

Neither is this coordination problem likely to be solved by granting some type of property right in the individual allelic information. Often granting a property right in a commodity will encourage private exchange, allowing markets to form – indeed, this is a large part of the impetus behind granting patents. However, valuing aggregated individualized data presents a problem that is well known to economic analysis of personal privacy.⁴⁷ Personalized data, such as that relating to consumer behavior, demographics, or other individual characteristics, is of relatively little value unless aggregated.⁴⁸ Because the value of any individual’s data is overshadowed by potential transaction costs, markets for such data do not emerge: no individual datum, taken in isolation, is valuable enough to bargaining over it worthwhile. Not surprisingly, consumers take very little care to guard such information, although they may feel some sense of entitlement or disenfranchisement when they discover their data has become part of a larger aggregation.⁴⁹

In essence there remains a disjunction between the private and public valuation of personal data. Although proposals have been made to treat individual data as property,⁵⁰ the impediments to bargaining suggest there is little likelihood that such mechanisms will align public and private

⁴⁶ The “Prisoner’s Dilemma” model serves as a standard example of a non-cooperative game in which the rational self-interest of the players leads to a suboptimal outcome for each player. DAVID M. KREPS, *GAME THEORY AND ECONOMIC MODELING* 37-39 (1990). In at least some instances, however, this dynamic may change if the game continues through multiple rounds. See ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* (1984).

⁴⁷ See Doreen Starke-Meyering, Dan L. Burk & Laura J. Gurak, *American Internet Users and Privacy: A Safe Harbor of Their Own?*, in *SOCIETY ONLINE: THE INTERNET IN CONTEXT* 275 (Philip E.N. Howard & Steve Jones eds., 2003).

⁴⁸ In this regard, personal data in a database shows characteristics of economic network effects albeit on the supply side rather than the demand side; it becomes more valuable as it is associated with other data. See generally Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479 (1998) (describing network effects).

⁴⁹ See Starke-Meyering, Burk & Gurak, *supra* note 47.

⁵⁰ See, e.g., Ann Bartow, *Our Data, Ourselves: Privacy, Propertization, and Gender*, 34 U.S.F. L. REV. 633, 634 (2000); Richard S. Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L.J. 2381, 2384 (1996); Catherine M. Valerio Barrad, *Genetic Information and Property Theory*, 87 NW. U. L. REV. 1037, 1039 (1993).

incentives.⁵¹ And, as with other personal data, the private exchange value of individual diagnostic testing information is typically negligible, as opposed to the public value of testing information. No individual datum or personalized set of allelic data is worth very much; only when assembled into large sets that reveal population frequencies does the conglomerated data become valuable. Consequently, a property based solution such as granting an exclusive right to the individuals from whom the data is derived is unlikely to create a market in the data; the transaction costs of negotiating and licensing and given individual's data will overwhelm the value of the data exchange itself, making the exchange unprofitable.

At the same time, the corollary is that no particular patient has any leverage sufficient to stifle the entire project. Unlike "holdout" scenarios where the permission of multiple rights holders is essential,⁵² and any given rights holder can bargain for a price that attempts to capture the value of the entire project, no particular piece of aggregated data is sufficiently necessary that its absence threatens the functionality of the completed set. Value in statistically large data sets derives from the correlations between individual entries, and in a large population, any given entry is largely fungible and can be replaced. If a particular individual does not supply the sequence of a given allele, it can be equally well obtained from others in the population, and the correlation of that allele to a given phenotype derives from the aggregation of multiple instances, and is not dependent on the inclusion of any given instance.

PATENTS AS AGGREGATORS

The economic configuration that I have just described for test data poses severe impediments to creation of an allelic database. And yet one exists – the one held by Myriad. This suggests that the analysis of gene patenting involves a more complicated story than the benefit or detriment to consumer pricing as a result of patenting, and for that matter more complicated than the benefit or detriment to investment in diagnostic testing that might result from patent exclusivity.

One important but previously overlooked part of this story may be the value of a patent as a coordination point. Myriad was successful in gathering in a single database a large body of information regarding BRCA1 and BRCA2 alleles due to the monopoly position conferred by the patent. Because Myriad was the only test provider to whom samples could be sent, information about the samples that might otherwise have been dispersed among different non-cooperating test providers was collected at a single location. The information gathered was not necessarily part of the invention covered by the patent; rather,

⁵¹ Pamela Samuelson, *Privacy as Intellectual Property?*, 52 STAN L. REV. 1125 (2000).

⁵² See MICHAEL HELLER, *THE GRIDLOCK ECONOMY: HOW TOO MUCH OWNERSHIP WRECKS MARKETS, STOPS INNOVATION, AND COSTS LIVES* (2008).

the patent created an exclusive condition, resulting in aggregation of associated data at a single locus.

The downside of this story, of course, is that the know-how collected during the lifetime of the patent remains proprietary with the patent holder. This suggests that patents may serve to aggregate dispersed information, but the aggregated data may not benefit the public, or may do so in only a limited way. In this particular case, Myriad leveraged patents of a type no longer available, which have now been declared invalid by the Supreme Court. But whether the focal technology remains patentable subject matter or not, all patents expire eventually, and their claimed inventions pass into the public domain. Associated, undisclosed know-how does not necessarily pass into general accessibility along with the patented technology to which it relates, it may rather be retained as a trade secret as Myriad appears to be doing with its accumulated data on BRCA1 and BRCA2 alleles.

The BRCA1 and BRCA2 data aggregation scenario provides an example of patent leverage in the context of a particular, highly publicized, and controversial diagnostic test. However, there is no reason to believe that this scenario is limited to polymorphisms in nucleotide sequences, let alone these particular polymorphisms as patented by Myriad. Quite the contrary: many other macromolecules, which may be the cause of disease or the target of pharmaceutical action, display patterns of microheterogeneity that will lend themselves to diagnosis or treatment only when information about their variations is collected and collated.⁵³ It would be equally challenging to aggregate information on, say, particular patterns of enzyme glycosylation⁵⁴ and their correlation to certain diseases, or on the correlation between particular patterns of enzyme glycosylation and the efficacy of particular enzymatic inhibitors.⁵⁵ The problem of dispersed information highlighted by the Myriad database scenario is likely endemic to the “omics” revolution; as genomics gives way to proteomics and lipidomics and metabolomics and kitchensinkomics,⁵⁶ aggregation of multiple data sets correlating individual variations with clinical manifestations will be important to understanding the overall pattern of disease.

⁵³ See Holman, *supra* note 10, at 645.

⁵⁴ The attachment of carbohydrate side chains to proteins is termed glycosylation. See ROSLYN M. BILL, LEIGH REVERS & IAIN WILSON, *PROTEIN GLYCOSYLATION* 8 (1998).

⁵⁵ See Douglas McCormick, *Healthcare* in *BIOTECHNOLOGY: THE SCIENCE AND THE BUSINESS* 387, 545-55 (Derek G. Springham, Vivian Moses, & Ronald E. Cape eds., 2d ed. 1999) (discussing correlation of glycosylation microheterogeneity to disease); Hook, *supra* note 18.

⁵⁶ See H. Ge, A.J. Walhout, & M. Vidal, *Integrating “omic” information: a bridge between genomics and systems biology*, 19 *TRENDS IN GENETICS* 551 (2003) (discussing the “omics” revolution); *Proteomics, transcriptomics: what’s in a name?*, 402 *NATURE* 715, 716 (1999) (describing extensions of the “omics” nomenclature).

Taking the BRCA1 gene as an example, one could imagine that microheterogeneity at several different sites, and different levels of cellular operation, might contribute to cancer susceptibility. A variety of small mutations in the genetic sequence, to which the Myriad patents were directed, might incapacitate the protective protein. But variations in glycosylation or other post-translational modification of the protein might also render it inoperative or less effective. Idiosyncratic methylation⁵⁷ of the gene itself, whether mutated or unmutated, might also inhibit production of the protein. Similarly, alterations in one or more of the other proteins that function with BRCA1 in a large complex might render the protein ineffective. Any one of these might inhibit function of the BRCA1 protein, or the effect may be cumulative; an impaired protein with reduced activity might still be effective if produced in sufficient quantities, or if unencumbered by additional impairments at other stages of its operation.

One might expect information on any or all of these possible factors to be collected at the point where genetic testing of the BRCA1 gene occurred, that is, where the patent was practiced or authorized. Patents that claimed novel nucleotide sequences would also logically include RNA transcripts, peptides, anti-sense sequences, and other molecules sharing the same sequence basis – related claims for different embodiments of the same invention. Diagnostics for each of these molecules might tend to coalesce around firms holding such patents. But such predictable structural patterns are precisely what the Supreme Court excluded from patentable subject matter in the first half of its *Myriad* opinion.⁵⁸ Thus, aggregation of such related data around such patents seems no longer a viable option since the decision.

The aftermath of the *Myriad* patent decision thus suggests the possibility that diagnostic innovation could move to a model in which proprietary data, or trade secrets, form the basis for competitive value in personalized medicine. Of course, patented technologies routinely accumulate associated trade secrets. Development or operation of a patented industrial process, or a patented machine, will often generate know-how about the quirks of the invention: about how it is most efficiently managed or maintained, about which customers are most interested in its benefits, about where to place or deploy it in a factory or other production space, and a thousand other details regarding its practical operation. Some of this information may accrue after the filing of the patent application, and so will not appear in the patent. Other information may be known at the time of filing, but unless it is necessary to making and

⁵⁷ Chemical addition of methyl groups to DNA bases is one form of gene regulation in normal cells. See Theresa Phillips, *The Role of Methylation in Gene Expression*, 1 NATURE EDUCATION 116 (2008), available at <http://www.nature.com/wls/topicpage/The-Role-of-Methylation-in-Gene-Expression-1070> (archived at <http://perma.cc/3V27-NKZQ>).

⁵⁸ See Dan L. Burk, *The Curious Incident of the Supreme Court in Myriad Genetics*, 90 NOTRE DAME L. REV. 505, 509 (2014).

2015]

PATENTS AS DATA AGGREGATORS

using the claimed invention, may well be excluded from the disclosure required to obtain the patent. And some information, such as the customer list associated with the invention, may be extremely valuable from a business standpoint but simply is not germane to the technical patent disclosure at all.

This view of patents as information collection points is largely coherent with general observations regarding the relationship between patents and know-how in bargaining situations. A number of commentators have observed that what is disclosed in a patent is likely only a small portion of the knowledge associated with a particular invention.⁵⁹ We have already noted that some information associated with an invention may be know-how that is not required by the patent disclosure but which may still be useful in practicing the invention. Other, associated knowledge may be tacit knowledge that is difficult or impossible to codify and convey in the text of a patent.⁶⁰ Such tacit knowledge will instead be carried by skilled individuals and become accessible to a licensee only once the proper individuals are identified; thus the patent becomes in essence a signal that a firm has employees with valuable related knowledge.⁶¹

Often such associated knowledge may be more valuable than the information disclosed in the patent itself. In such cases, what a licensee desires from a patent license may actually be access to the know-how available from the patent holder; frequently this means access to the patent holder's personnel who are skilled in implementing the patented technology.⁶² This suggests that patents provide only the thin edge of the licensing wedge, the tip of an informational iceberg, adding value not so much due to their own exclusivity, nor even due to their role in disclosing how to make and use the claimed invention, but as due to their role in disclosing the nature of the technological expertise held by the licensor.

The same is doubtless true of diagnostic methods that may be the subject of patents; there is likely a considerable body of proprietary know-how associated with implementing such an invention. But the focus of my argument here is somewhat different; I am concerned not with trade secrets that are ancillary to the patent's subject matter, but with trade secrets that are *consequent to the patent's exclusivity*. Exclusivity is not germane to the routine accumulation of know-how around the invention – any invention, even an unpatentable invention, may generate associated know-how; if the unpatentable invention is kept as a trade secret, it may beget further trade secrets.⁶³ Here I am focusing

⁵⁹ Ashish Arora, *Contracting for Tacit Knowledge: The Provision of Technical Services in Technology Licensing Contracts*, 50 J. DEV. ECON. 233, 246 (1996).

⁶⁰ See Dan L. Burk, *The Role of Patent Law in Knowledge Codification*, 23 BERKELEY TECH L.J. 1009 (2008).

⁶¹ See Arora, *supra* note 59, at 252.

⁶² *Id.*

⁶³ Thus, Sichelman & Simon suggest that search engines or social media may offer a

not on the inevitable array of know-how surrounding the implementation of the claimed invention, but instead on the effect of patent exclusivity on the aggregation of already existing but widely dispersed information. The allelic database held by Myriad exists as a result of aggregation under patent exclusivity, not merely due to the accumulation of tacit knowledge incident to performing the diagnostic process.

TOWARD AN AGGREGATION THEORY

The new Myriad strategy of trade secrecy, and the potential for patent exclusivity to serve as an aggregation point for dispersed data, should probably come as no surprise given the legal structure of these forms of intellectual property and our understanding of the economics of interaction between the two. As a legal matter, trade secrets convey no right to exclude, only a right against misappropriation.⁶⁴ The proprietary information entailed in a trade secret can legitimately be either reverse engineered or independently re-created.⁶⁵ This tends to curb the tendency toward market power; since the trade secret carries no exclusivity against the world, licensees of a trade secret are only bargaining for disclosure, and the legally permissible alternatives to licensing build potential substitutes into the system. A trade secret owner who tries to price licenses at a cost higher than the cost of independent development or reverse engineering the secret will see competitors adopt the lower-cost option.⁶⁶

Transfer of trade secrets is therefore fraught with the problem known as Arrow's Information Disclosure Paradox: once information is disclosed it is simultaneously held by both parties to the transaction, which may tend to deter disclosures.⁶⁷ Unlike licensing or exchange of a physical artifact, there is no natural barrier, such as possession, to prevent a party from behaving strategically by walking away with the newly-disclosed information. Consequently, knowing that information once disclosed cannot be recalled, there will be reluctance to disclose such information absent some assurance of

profile similar to diagnostic tests for proprietary data aggregation. *See* Ted Sichelman & Barbara Simon, *Generating Trade Secrets From Patents*, Presentation at the 14th Annual Intellectual Property Scholar's Conference, Berkeley, California (Aug. 7, 2014). However, even if the search algorithm or media platform is patented, data aggregates around search engine operation by virtue of the search function, and around social media as a result of user subscription, not as a result of legal exclusivity.

⁶⁴ RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 43 (1995).

⁶⁵ *Id.*; UNIF. TRADE SECRETS ACT § 1, cmt; RESTATEMENT OF TORTS, § 757 CMT. B (1939).

⁶⁶ *See* Martin J. Adelman, *Property Rights Theory and Patent-Antitrust: The Role of Compulsory Licensing*, 52 N.Y.U. L. REV. 977, 981 (1977).

⁶⁷ *See* Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY* 609 (Nat'l Bureau of Econ. Research ed., 1962).

confidentiality. However, the potential recipient of the information will be simultaneously reluctant to commit to confidentiality without knowing the nature of the commitment. But knowing the nature of the commitment would of course require disclosure, recursively bringing the problem back to the same beginning. Additionally, to the extent that the parties may attempt to avoid this problem by using contract law to forestall strategic behavior after disclosure, contracts are necessarily incomplete, and will be especially vague in advance of actual disclosure, potentially leaving considerable latitude for strategic behavior.⁶⁸ Thus the parties attempting to bargain over trade secrets may reach an impasse regarding disclosure, and valuable collaborations may be deterred.

Edmund Kitch long ago observed that patents offer a solution to the Information Disclosure Paradox, overcoming the barrier to licensing valuable information.⁶⁹ By publicly disclosing valuable information while subjecting it to a strong right of exclusion, patents allow parties negotiating over technology to know what is being offered for licensing without either the concern that it will be misappropriated when disclosed, or the costs of elaborate contractual maneuvering to avoid such misappropriation.

Kitch also suggested that patents may serve a coordination function, which presages the aggregation function I have discussed here. The Myriad database story implicates what has come to be known as the “prospect theory” of patents, but with a bit of a twist not typically seen in past patent prospect apoloia. Kitch famously proposed a rationale justifying patents on the basis of a comparison to physical resources, specifically comparing intellectual property rights to mining claims.⁷⁰ As a result of the analogy to mineral prospecting, his explanation of patent benefits became known as the “prospect” theory. Kitch proposed that, just as granting rights in mineral prospects or other valuable resources facilitates the orderly and coordinated exploitation of the resource, so patents serve to vest rights to a new technology in a single overseer who can coordinate development of the technology.⁷¹

Kitch was specifically interested in tying the rationale for intellectual property to the broader theory of property rights, particularly that formulated by Demsetz, which holds that private ownership creates a self-interested incentive in metering the exploitation of an exhaustible resource.⁷² Private ownership attempts to mediate private and public valuation of a resource, so as to avoid the “tragedy of the commons” which might occur when an unmanaged

⁶⁸ Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 278 (1977).

⁶⁹ *Id.* at 266, 277–78.

⁷⁰ *See id.*

⁷¹ *Id.* at 270-71.

⁷² *See* Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347 (1967); *see also* Brett Frischman, 3 REV. L. ECON. 649 (2007) (questioning the applicability of Demsetz’s approach to intellectual property).

physical resource is exhausted through over-exploitation.⁷³ Property rights in physical resources help align public and private values.⁷⁴ Kitch similarly hoped to develop a theory of management for innovation resources, arguing that the coordination provided by a patent could prevent wasteful dissipation of an invention's value that might occur through duplicative or haphazard development.⁷⁵

Although widely discussed,⁷⁶ the prospect approach has not been widely accepted, in part because detractors have argued that unlike physical resources, intellectual property is non-rivalrous and essentially inexhaustible, so that there is no likelihood of a tragedy of the commons, and no finite resource to coordinate.⁷⁷ What is of primary interest in the context of personalized medicine is not so much that question, but rather the connection to another of Kitch's pioneering intellectual property analyses, concerning the development of unpatented but valuable information.⁷⁸ I have pointed out above the disparity in private and public valuation of individual data, and the difficulty in aligning such value.⁷⁹ In the situation I have outlined for BRCA1 and BRCA2 alleles, the (now extinguished) patent is not so much the tool that allows a single entrepreneur to coordinate development of the patented technology itself as it is the tool that allows coordination of associated know-how. If patents are to be compared to mineral rights, the benefit in this case is perhaps not the facility of the property right in coordinating the orderly extraction and processing of the ore, so much as it is the ability of the property right to organize placement of outbuildings and systematic scheduling of the crew around the mine.

⁷³ See Garrett Hardin, *The Tragedy of the Commons*, 162 SCI. 1243 (1968).

⁷⁴ See Demsetz, *supra* note 72, at 356.

⁷⁵ See Kitch, *supra* note 68, at 276-78; see also Wendy J. Gordon, *Asymmetric Market Failure and Prisoner's Dilemma in Intellectual Property*, 17 U. DAYTON L. REV. 853, 855-56 n.13 (1992) (applying "coordination" rationale to copyright). Robert Denicola has also made a similar argument in the context of trademarks, which as signals rather than as discrete goods, are not generally included in the "incentive" rationale of patent and copyright. See Robert Denicola, *Institutional Publicity Rights: An Analysis of the Merchandising of Famous Trade Symbols*, 62 N.C. L. REV. 603, 637-41 (1984).

⁷⁶ See, e.g., Michael Abramowicz, *The Danger of Underdeveloped Patent Prospects*, 92 CORNELL L. REV. 1065 (2007); John F. Duffy, *Rethinking the Prospect Theory of Patents*, 71 U. CHI. L. REV. 439 (2004).

⁷⁷ See Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 1048-72 (cataloging practical and conceptual shortcomings of the prospect theory); see also Mark Lemley, *Ex Ante Versus Ex Post Justifications for Intellectual Property*, 71 U. CHI. L. REV. 129, 139-40 (2004) (noting that Kitch's theory unnecessarily discounts market ordering).

⁷⁸ Edmund W. Kitch, *The Law and Economics of Rights in Valuable Information*, 9 J.L. STUD. 683 (1980).

⁷⁹ See *supra* notes 47-51 and accompanying text.

2015]

PATENTS AS DATA AGGREGATORS

But this coordination of non-patent data plays havoc with another of Kitch's arguments, which addresses patent rights and market power. In a separate work addressing the market power of patents, Kitch argued that patents are unlikely to allow monopolistic pricing, in part because the patent has a finite lifespan, and in part because pricing as a monopolist excludes low-valuing customers and so obscures their identity.⁸⁰ Because the patent is subject to a definite endpoint, patent holders will be aware of the temporal limit of their exclusivity, and will wish to identify the largest possible swath of potential customers during the patent term, so that those customers can be cultivated and recruited to brand loyalty before the exclusivity expires and other competitors enter the field.⁸¹ Monopoly pricing prevents the patent holder from finding and cultivating such customers, and so would rationally be avoided during the patent term.

The Myriad aggregation scenario places the patent holder in a very different position. Aggregation of proprietary information essentially allows the patent holder to extend aspects of exclusivity beyond the term of the patent. Patent exclusivity results in data aggregation and the information aggregated around the patent may remain proprietary indefinitely as trade secrets, depending on the cost of independent discovery or re-creation. As previously suggested, for aggregated data such as genomic variants, the transaction costs of re-creating the database may be prohibitive, so that no market substitutes become available. Taking the contrapositive of the Kitchian anti-monopoly argument, this result suggests not only that the patent holder may be able to leverage the aggregated information into elevated pricing beyond the term of the patent, but knowing that the expiration of the patent may also have less incentive to avoid monopoly pricing during the term of the patent itself.

PROMOTING AGGREGATION AND ACCESS

What the *Myriad* situation illustrates, and what the analysis above is intended to unpack, is a fairly difficult dilemma requiring a choice between two possible related but distinct and undesirable states of the world: the first in which widely dispersed information is aggregated due to the exclusivity of a patent, but is unavailable to the public due to trade secrecy, and the second where widely dispersed information is never aggregated in the first place, because no patent is available as a focus. In the first state social welfare suffers because a valuable resource may be available only on a limited and costly basis at the behest of a trade secret owner, and in the second social welfare suffers because the valuable resource is not available at all. Resolving the former problem involves overcoming a monopolistic market failure, resolving the

⁸⁰ Edmund W. Kitch, *Patents: Monopolies or Property Rights?*, 8 RES. L. & ECON. 31 (1986).

⁸¹ *Id.*

second involves overcoming a collective action transaction costs market failure. Patents may resolve the second problem by creating the first; removing patents resolves the first problem by creating the second.

This is a knotty and possibly intractable set of problems, which I suggest has probably been worsened by the Supreme Court's elimination of potentially aggregating patents — not to mention the concomitant uncertainty surrounding the question of whether and to what extent the Supreme Court really *has* eliminated potentially aggregating patents.⁸² In the presence of patents, social welfare may be somewhat farther ahead, as the disparate information related to the claimed invention becomes concentrated in a single location. Society is likely better off having the database than not having it. Scattered, the information has little or no value; gathered, it becomes a useful and valuable artifact that might not have come into existence without patent exclusivity to direct it to a common overseer. The downside, of course, is that it remains in the hands of a single user who is able to behave as a monopolist. While this is better than not aggregating the data at all, it may be inferior to aggregating the data in a manner that either is, or becomes, more widely accessible.

Given that the Supreme Court may have eliminated patents as aggregators in a variety of situations, a different aggregation mechanism could be needed. One could imagine drastic centralization regimes, for example where all genetic testing was done by a centralized government facility, such as the National Institutes of Health, or by licensed government labs, so that the aggregated data was available as a public resource. Or all newly discovered instances of microheterogeneity might have to be reported to a government source, such as the Centers for Disease Control.⁸³ In such systems, governmental fiat replaces patents as the point of aggregation. The resource comes into existence at the point of regulatory control. However, the benefits of private innovation and competition would be lost, and the latter scheme might well implicate constitutional takings.

A system harnessing private ordering would likely be preferable, if the differential between public and private value can be reconciled. Fortunately, the problem of aggregating disparate proprietary data is not entirely unique to the diagnostic area and some work has already been done to begin addressing it in related contexts. Copyright law has famously produced licensing societies to collectively administer the public performance right for music composers, in situations such as broadcast where the transaction costs of individual negotiations would be prohibitive.⁸⁴ For access to a portfolio of compositions, licensees pay a single fee which is then divided and administered by the

⁸² See Burk, *supra* note 58, at 508-09 (pondering the extent of the *Myriad* holding).

⁸³ See, e.g., Cook-Degan et al., *supra* note 41, at 587 (arguing for systems of government mandated data disclosure).

⁸⁴ See Robert P. Merges, *Contracting into Liability Rules: Intellectual Property Rights and Collective Rights Organizations*, 84 CALIF. L. REV. 1293 (1996).

licensing society. One can imagine a similar structure, in which the holders of allelic or other relevant data contribute their information under a collective trade secrecy agreement, and in return gain access to others' data and perhaps some pro rata payment each time the database is used.⁸⁵ It is worth noting, however, that copyright licensing societies exist only under sufferance of antitrust consent decrees, so that some accommodation under competition laws might be necessary.⁸⁶

Of course, unlike musical composition copyrights, trade secrets generally bear the hallmarks of liability rules rather than property rules: trade secrets are efficacious only against misappropriation by select actors in contractual or fiduciary relationships to the trade secret holder; they are not good against the world.⁸⁷ Classic work by Professor Merges suggests that information pooling mechanisms work better in the presence of property rules, such as patents or copyrights, which allow the holders to "contract into liability rules" if they choose.⁸⁸ But as Professor Lemley has recently pointed out, the converse is clearly also true: holders of liability rules can and do contract into property regimes, suggesting that private ordering may be able to overcome the pooling problem for trade secrets.⁸⁹

The copyright collecting society has previously been suggested as a model in situations related to the diagnostic situation. A version of the dispersed proprietary data problem has long been endemic in small-molecule drug development, although presumably not to the same degree for mass market pharmaceuticals that it may present a problem for future development of personalized treatments.⁹⁰ Suites of information about newly developed pharmaceutical candidate molecules tend to exist separately from suites of information about potential targets of action, and the tendency to hold such data portfolios separately as trade secrets presents a severe impediment to drug discovery and development.⁹¹ Rai et al. have argued that the drug discovery problem might be addressed by construction of a contractual semi-commons,

⁸⁵ This would be similar in several respects to proposals by Rai et al. for pharmaceutical development. See Arti K. Rai et al., *Pathways Across the Valley of Death: Novel Intellectual Property Strategies for Accelerated Drug Discovery*, 8 YALE J. HEALTH POL'Y L. & ETHICS 1 (2008).

⁸⁶ See Julie E. Cohen, *Lochner in Cyberspace: The New Economic Orthodoxy of "Rights Management"*, 97 MICH. L. REV. 462, 500-01 (1998).

⁸⁷ See David Friedman et al., *Some Economics of Trade Secret Law*, 5 J. ECON. PERSP. 61, 62 (1991) (defining trade secrecy as a liability rule regime). But see Mark A. Lemley, *The Surprising Virtues of Treating Trade Secrets as IP Rights*, 61 STAN. L. REV. 311, 313 (2008) (arguing that trade secrets are best understood as a species of intellectual property).

⁸⁸ See Merges, *supra* note 84.

⁸⁹ Mark A. Lemley, *Contracting Around Liability Rules*, 100 CALIF. L. REV. 463 (2012).

⁹⁰ See Rai et al., *supra* note 85 at 4.

⁹¹ *Id.* at 3-4.

into which data owners could release proprietary data, subject to confidentiality agreements. Candidate molecules and targets could be tested against one another, and the results held in escrow by a third party. Contributors to the testing data pool would be pro rata compensated if products were developed using testing data derived from their contributions. One can imagine a similar model might be adopted to facilitate diagnostic data licensing.

Alternatively, in areas that the *Myriad* opinion perhaps does not reach, patents may still be available as an aggregation strategy, perhaps in some instances on a limited basis. For a variety of reasons, this state of affairs appears somewhat less problematic than the situation without patents. First, as I have indicated above, under a regime of exclusivity, the data will become aggregated and the public will gain the benefit, although unfortunately not at competitive pricing. Second, the patent version of the problem, for better or for worse, is temporary. Patents expire in about twenty years,⁹² at which point the problem converts to a problem without exclusivity. But it does not convert to the problem without patents as discussed above; rather, the patent version of the problem converts to a problem of concentration in the hands of one former patent holder, rather than a problem of inchoate dispersion.

And third, this concentration problem is a familiar problem that probably does not require any new tools or elaborate pooling mechanisms to address. As I have outlined above, trade secret protection of proprietary data that has been aggregated under patent exclusivity may be viewed as effectively an indefinite extension of the benefits of the patent. Federal patent precedent already addresses some types of situations in which patent holders have attempted to use ancillary state laws to leverage the benefits of their patent beyond the term set for exclusivity.

For example, the Supreme Court has in the past reprimanded patent holders attempting to extend their period of patent exclusivity by subsequently asserting trademark brand recognition that accrues in their patented product during the exclusive period.⁹³ Under other circumstances, the brand would legitimately merit trademark protection, but the Court has made clear that federal patent policy allows only a defined period of exclusivity, which cannot be effectively extended using trademark law; consequently trademark protection in those situations is extinguished.⁹⁴ Similarly, the Court has made clear that state contract law cannot be leveraged to effectively extend the period of patent exclusivity; federal policy preempts state law alternatives that interfere with the duration of protection set by Congress.⁹⁵

⁹² 35 U.S.C. § 154(a)(2) (2012).

⁹³ See, e.g., *TrafFix Devices, Inc. v. Mktg. Displays, Inc.*, 532 U.S. 23, 29-30 (2001); *Kellogg Co. v. Nat'l Biscuit Co.*, 305 U.S. 111, 122 (1938).

⁹⁴ See Dan L. Burk, *Cybermarks*, 94 MINN. L. REV. 1375, 1407 (2007).

⁹⁵ See *Brulotte v. Thys Co.*, 379 U.S. 29, 32 (1964).

One can imagine a comparable rationale preempting state trade secrecy protection where it is leveraged to extend the benefits of the patent grant beyond the end of the patent. Admittedly, the Supreme Court has squarely held in *Kewanee Oil v. Bicron Corp.* that state trade secrecy and federal patents can peacefully co-exist,⁹⁶ but this by no means necessarily insulates trade secrecy from particularized preemption.⁹⁷ Trade secrecy co-exists with patent law as a general matter, just as state trademark and contract law do. It is clear from the Supreme Court's jurisprudence that although these latter bodies of law co-exist with patent law, this does not preclude conflict preemption of particular trademarks and contracts in situations where they impede federal policy.⁹⁸ Particular instances of trade secrecy that impede federal policy would be logically be treated the same.⁹⁹

Preemption of trade secrets of course could prove messy; one would not want to open the flood gates to a run of employee data pilfering or industrial espionage, prompted by the prospect of hiding behind federal supremacy. Consequently, it may sometimes be preferable to employ alternative tools from the same jurisprudence on federal patent primacy. In some cases, rather than being preempted by federal law, state law arrangements to extend the benefits of the patent past the prescribed term have been held to constitute patent misuse.¹⁰⁰ Misuse is an equitable doctrine related to unclean hands, in which courts refuse to enforce a patent that has been improperly leveraged beyond the exclusivity intended in the patent grant.¹⁰¹ Refusal to enforce the patent would of course only be helpful to remedy monopolistic aggregation during the life of the patent, and not so much of a threat after the patent's expiration. But the doctrine has been extended to other intellectual property rights, such as copyright, when overextended via state contract law.¹⁰² There is no particular

⁹⁶ 416 U.S. 470, 493 (1974).

⁹⁷ See Dan L. Burk, *Misappropriation of Trade Secrets in Biotechnology Licensing*, 4 ALB. L.J. SCI. & TECH. 121, 148 (1994).

⁹⁸ *TraFFix Devices*, 532 U.S. at 33 (finding that federal patent policy precludes trade dress protection of functional though distinctive product design); *Lear, Inc. v. Adkins*, 395 U.S. 653, 673 (1969) (deciding that patent policy precludes contractual obligations not to challenge patent validity). *But see also* *Aronson v. Quick Point Pencil Co.*, 440 U.S. 257, 263-64 (1979) (upholding properly constructed licensing of trade secret disclosure after failed patent application).

⁹⁹ Burk, *supra* note 94.

¹⁰⁰ See *Ill. Tool Works Inc. v. Indep. Ink, Inc.*, 547 U.S. 28, 46 (2006) (finding that patent misuse requires that a practice "has the effect of extending the patentee's statutory rights . . . with an anti-competitive effect"); *Windsurfing Int'l, Inc. v. AMF, Inc.*, 782 F.2d 995, 1001 (Fed. Cir. 1986) (explaining that misuse entails impermissibly broadening the "physical or temporal scope" of a patent).

¹⁰¹ See Thomas F. Cotter, *Misuse*, 44 HOUS. L. REV. 901, 901-02 (2007).

¹⁰² See, e.g., *Alcatel USA, Inc. v. DGI Techs., Inc.*, 166 F.3d 772, 792-93 (5th Cir. 1999); *Practice Mgmt. Info. Corp. v. Am. Med. Ass'n*, 121 F.3d 516, 520-21 (9th Cir.

reason why a court might not apply the doctrine to trade secrets as a sort of “fruit of the poisonous tree” exclusion for leveraging patent benefits beyond the relevant patent term.

As a third alternative, anticompetitive positioning by an intellectual property holder might be addressed via antitrust law proper. Current antitrust standards tend to discount the market power afforded by trade secrets,¹⁰³ since as noted above independent recreation or reverse engineering offer the possibility of non-infringing substitutes; hence the ability to exclude by withholding disclosure is limited, and can only be effective up to the point where permissible re-creation of the secret costs less.¹⁰⁴ However, antitrust violation by means of trade secrecy is not impossible. If the information is costly or impossible to re-create, there exists the possibility that it might confer sufficient market power to constitute monopolization or an attempt to monopolize. The anti-trust constraint might additionally bear consideration in light of the Kitchian argument against patent monopoly, reviewed above.¹⁰⁵ If Kitch is correct that the certain end of the patent term deters monopoly pricing, then the prospect of proprietary data aggregation may negate the incentive against monopoly pricing during the term of the patent, by virtue of effectively eliminating the sunset horizon of the patent period. If no competitors are able to enter the market at the end of the patent period, monopoly pricing might be maintained during the patent period and beyond, inviting antitrust scrutiny.

To be sure, all of these latter approaches rely on older case law that is currently disfavored by the Federal Circuit. The Federal Circuit has over time greatly restricted the application of the misuse doctrine,¹⁰⁶ and has shown little enthusiasm for applying competition law to curb expansive uses of patents.¹⁰⁷ But in a recent challenge to the per se prohibition on post-termination patent licensing, the Supreme Court has reaffirmed the viability of this doctrine, giving it a new lease on life.¹⁰⁸ Despite the likely reluctance of the Federal

1997); *Lasercomb Am., Inc. v. Reynolds*, 911 F.2d 970, 976-77 (4th Cir. 1990); *see also* Dan L. Burk, *Anti-Circumvention Misuse*, 50 *UCLA L. REV.* 1095 (2003) (arguing for further extension of misuse doctrine to DRM anti-circumvention privileges).

¹⁰³ *See* Harry First, *Trade Secrets and Antitrust Law*, in *THE LAW AND THEORY OF TRADE SECRECY: A HANDBOOK OF CONTEMPORARY RESEARCH* (Rochelle C. Dreyfuss & Katherine J. Strandburg eds., 2011).

¹⁰⁴ *See* Adelman, *supra* note 66 and accompanying text.

¹⁰⁵ *See supra* note 80 and accompanying text.

¹⁰⁶ *USM Corp. v. SPS Techs., Inc.*, 694 F.2d 505, 510 (7th Cir. 1982) (opining that misuse is “confined to a handful of specific practices”). Cf. *Princo Corp. v. Int’l Trade Comm’n*, 616 F.3d 1318, 1342 (Fed. Cir. 2010) (Dyk, J., dissenting) (arguing against an en banc majority that would “emasculate” or entirely eliminate the doctrine).

¹⁰⁷ DARYL LIM, *PATENT MISUSE AND ANTITRUST LAW: EMPIRICAL, DOCTRINAL, AND POLICY PERSPECTIVES* 63-65 (2013).

¹⁰⁸ *Kimble v. Marvel Entm’t., Inc.*, 135 S. Ct. 2401 (2014).

2015]

PATENTS AS DATA AGGREGATORS

Circuit to employ misuse and related measures, reinvigorating these doctrines requires simply the will to use existing tools, rather than the necessity of creating new ones.

CONCLUSION

The role of patents as data aggregators to some degree resembles the old vaudeville routine saying about having some good news and having some bad news.¹⁰⁹ The good news is that the exclusive effects of a patent may provide a focal point for aggregation of widely dispersed but potentially valuable data that might otherwise remain dispersed and unusable. The bad news is that the aggregated data may then become trade secrets or otherwise be captured¹¹⁰ as proprietary know-how to the patent holder. Given that no one else in the market will have the exclusivity lever of the patent to aggregate competing data, the information may remain proprietary even after the focal patent, around which the data was collected, is long since expired.

Both the good news and the bad news have become salient in the wake of the *Myriad* decision. It may take some time to determine the full effect of the *Myriad* opinion, but it seems clear that patents may be unavailable for a variety of diagnostics and personalized treatments, either due to the *Myriad* subject matter restrictions or due to the preclusive effect of other patenting criteria. As I have suggested here, the problem for personalized medicine is whether some alternative mechanism can be developed, either to facilitate open aggregation of information on microheterogeneities in the absence of a patent focal point, or to prompt disclosure of such information where patents are available, and serve as aggregators, but leave the information in an inaccessibly proprietary status. In either event, the aftermath of the *Myriad* case may showcase patents in a new and unexpected role not usually contemplated by theories that typically justify patents.

¹⁰⁹ JAMES A MATISOFF, *BELSSINGS, CURSES, HOPES, AND FEARS: PSYCHO-OSTENSIVE EXPRESSIONS IN YIDDISH* xiii (2d ed. 2000).

¹¹⁰ Trade secrecy is not the only route to retaining control of aggregated data. For example, I have not discussed here the use of contracts and legal prohibitions against circumventing technical encryption as tools to capture and retain control of such data. *See generally* Burk, *supra* note 97 (discussing misuse of anti-circumvention provisions). In some countries other than the U.S., *sui generis* database statutes might be used to retain control of captured data. *See generally* Jerome H. Reichman & Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 VAND. L. REV. 51 (1997) (discussing data compilation regimes).