JAMES LOWE, RHODRI LENG, GIL VIRY, MARK WONG, NIKI VERMEULEN AND MIGUEL GARCÍA-SANCHO*

The Bricolage of Pig Genomics

ABSTRACT

The history of genomic research on the pig (*Sus scrofa*)—as uncovered through archival research, oral histories, and the analysis of a quantitative dataset and co-authorship network—demonstrates the importance of two distinct genealogies. These consist of research programs focused on agriculturally oriented genetics, on the one hand, and systematics research concerned with evolution and diversity, on the other. The relative weight of these two modes of research shifted following the production of a reference genome for the species from 2006 to 2011. Before this inflection point, the research captured in our networks mainly involved *intensive* sequencing that concentrated primarily on increasing the resolution of genomic data both in particular regions and more widely across the genome. Sequencing practices later became more *extensive*, with greater focus on the generation and comparison of sequence data across and between populations. We explain these shifts in research modes as a function of the availability, circulation, distribution, and exchange of genomic tools and resources—including data and materials—concerning the pig in general, and increasingly for particular populations. Consequently, we describe the history of pig genomics as constituting a kind of *bricolage*, in which geneticists cobbled together resources to which they had access—often ones produced by them for other purposes—in pursuit of their research aims. The concept of bricolage adds to the *thicker* vision of genomics that we have shown throughout the special issue and

*James Lowe, Science, Technology and Innovation Studies, University of Edinburgh, Old Surgeons' Hall, High School Yards, Edinburgh, EH1 1LZ, United Kingdom. jwelowe1985@gmail.com

The following abbreviations are used: ARS, Agricultural Research Service of the United States Department of Agriculture; BAC, Bacterial Artificial Chromosome; CEA, French Atomic Energy Commission; IMpRH, INRA-Minnesota porcine Radiation Hybrid panel; INRA, Institut National de La Recherche Agronomique; PiGMaP, Pig Gene Mapping Project; PMID, PubMed ID; SGSP, Swine Genome Sequencing Project; SLA, Swine Leucocyte Antigen system; UAB, Universitat Autònoma de Barcelona.

Historical Studies in the Natural Sciences, Vol. 52, Number 3, pps. 401–442. ISSN 1939-1811, electronic ISSN 1939-182X. DOI: https://doi.org/10.1525/hsns.2022.52.3.401.

further highlights the singularity of the dominant, thin narrative focused on the production of the human reference sequence at large-scale genome centers. This essay is part of a special issue entitled The Sequences and the Sequencers: A New Approach to Investigating the Emergence of Yeast, Human, and Pig Genomics, edited by Michael García-Sancho and James Lowe.

KEY WORDS: genomics, sequencing, mapping, genetics, *Sus scrofa*, pig, agriculture, systematics

1. HISTORIOGRAPHICAL BACKGROUND AND APPROACH

This paper provides new insights into the nature and organization of pig genomics. Pigs (*Sus scrofa*) are one of a select group of species to have been domesticated. In October 2021, the US Department of Agriculture estimated that 749,281,000 domesticated pigs were being kept as livestock in the nine largest pork-producing countries and the European Union.[1] Scientists use pigs as animal models in biomedical and surgical research, and are retooling them to better serve as a source of organs for transplantation into humans.[2] The salience of the relationships of pigs to humans, and the role of pork production culturally, socially, and economically, has encouraged a growing scholarly literature on these animals. This includes explorations of the ways in which pigs have enabled colonization as well as the development of industrialized capitalism and capitalist markets. The various political economic forces that have encouraged—and in turn been impacted by—the standardization of pigs within increasingly industrialized and large-scale models of pork production have also been examined. As well as economic and political impacts, these works have dealt with changing human relationships to pigs, and the consequences for farmers, other workers, and the wider environment of developments in pork production.[3] Historical

1. https://web.archive.org/web/20211209022748/https://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf

2. Eric M. Walters, Kevin D. Wells, Elizabeth C. Bryda, Susan Schommer, and Randall S. Prather, "Swine Models, Genomic Tools and Services to Enhance Our Understanding of Human Health and Diseases," *Lab Animal* 46 (2017): 167–72.

3. J. L. Anderson, *Capitalist Pigs: Pigs, Pork, and Power in America* (Morgantown: West Virginia University Press, 2018); Alex Blanchette, *Porkopolis: American Animality, Standardized Life, and the Factory Farm* (Durham, NC: Duke University Press, 2020); Thomas Fleischman, *Communist Pigs: An Animal History of East Germany's Rise and Fall* (Seattle: University of Washington Press, 2020); Robert Malcolmson and Stephanos Mastoris, *The English Pig: A History* (London: Hambledon Press, 1998); Tiago Saraiva, *Fascist Pigs: Technoscientific Organisms and the History of Fascism* (Cambridge, MA: MIT Press, 2018).

scholarship has also focused more directly on biological research concerning pig husbandry and production.[4]

Despite the importance of pig genetics and genomics to breeding and livestock production, and to the shaping of the pig as an animal model, little historical attention has been devoted to the mapping and sequencing of the pig genome, with the exception of reviews and accounts by the researchers themselves.[5] This paper contributes to the developing literature on the scientific, economic, and cultural importance of the pig, as well as demonstrating an agricultural motivation for research that complements our analysis of human and yeast genomics in this special issue. One of the main dimensions of agriculturally oriented genomics has been the use of DNA data to increase the accuracy and speed of selective breeding decisions in herds of pigs and other farm animal species: a recent instantiation of this is known as genomic selection.[6] The advent of genomic selection has aided the further concentration of the pig breeding industry, alongside the ongoing concentration of the production and processing sectors. Examining pig genetics and genomics allows us to inspect how a field of research deeply inflected with the problems

4. On reproductive science and embryology, see Paul Brassley, "Cutting across Nature? The History of Artificial Insemination in Pigs in the United Kingdom," *Studies in History and Philosophy of Biological and Biomedical Sciences* 38, no. 2 (2007): 442–61; Chris Polge, "The Work of the Animal Research Station, Cambridge," *Studies in History and Philosophy of Biological and Biomedical Sciences* 38, no. 2 (2007): 511–20. On science relating to pig health, see Abigail Woods, "Decentring Antibiotics: UK Responses to the Diseases of Intensive Pig Production (ca. 1925–65)," *Palgrave Communications*, 5 (2019): 41.

5. E.g. Bin Fan, Danielle M. Gorbach, and Max F. Rothschild, "The Pig Genome Project Has Plenty to Squeal about," *Cytogenetic and Genome Research* 134 (2011): 9–18; Martien A. M. Groenen, Lawrence B. Schook, and Alan L. Archibald, "Pig Genomics," in *The Genetics of the Pig*, 2nd edition, eds. Max F. Rothschild and Anatoly Ruvinsky (Wallingford, UK: CAB International, 2011), 179–99; Max F. Rothschild, "From a Sow's Ear to a Silk Purse: Real Progress in Porcine Genomics," *Cytogenetic and Genome Research* 102 (2003): 95–99. Works written outside the pig genetics and genomics community include Margaret Derry, *Masterminding Nature: The Breeding of Animals, 1750–2010* (Toronto: University of Toronto Press, 2015), 131–59, 185–86, and publications deriving from the same project in which the research for this paper was conducted: James W. E. Lowe, "Sequencing through Thick and Thin: Historiographical and Philosophical Implications," *Studies in History and Philosophy of Biological and Biomedical Sciences* 72 (2018): 10–27; James W. E. Lowe and Ann Bruce, "Genetics without genes? The Centrality of Genetic Markers in Livestock Genetics and Genomics," *History and Philosophy of the Life Sciences* 41: 50.

6. Lowe and Bruce, "Genetics" (n.5) situates genomic selection in the wider history of pig genetics and genomics. The technical details and commercial possibilities of genomic selection in pigs are assessed in Egbert F. Knol, Bjarne Nielsen, and Pieter W. Knap, "Genomic Selection in Commercial Pig Breeding," *Animal Frontiers* 6, no. 1 (2016): 15–22.

and interests of the breeding sector in turn shapes new developments in that industry through the production of data, knowledge, and various tools.

In genomics, as in other areas of science, accounts by scientists and the organization of their personal papers and other archival resources have encouraged historians to see research in terms of large-scale projects.[7] When embarking on our own project to map the history of pig genomics, studying particular projects was a tractable entry point. It is, however, an approach we have tried to avoid dominating our work so we can consider other kinds of collaborations that allow us to more directly identify the intended aims of genomic research.

In pig genomics, a series of collaborative projects emerged in the early 1990s, funded by (among others) the European Commission and the United States Department of Agriculture (USDA). These projects were conducted to identify and map genes and other markers to chromosomes to aid livestock improvement, for instance through selective breeding and transgenics.[8] The USDA's Agricultural Research Service (ARS) pursued an in-house effort, with factory-style mapping at the Meat Animal Research Center in Nebraska. Additionally, through the auspices of the Cooperative State Research, Education, and Extension Service, the USDA established the National Animal Genome Research Program, with a pig genome coordinator appointed to foster and support the burgeoning community of mappers and sequencers. The European Commission aided the development of collaborative networks around pig genomics and the sharing and circulation of materials and data through the funding of successive projects (such as the Pig Gene Mapping Project, PiGMaP, 1991–1996) and the concomitant blossoming of connections and capabilities across the continent.

The Swine Genome Sequencing Project (SGSP, 2006–2009) was a continuation of all these efforts. For this, the communities that had come together to map the pig genome in the 1990s contracted the Sanger Institute to sequence the pig genome and adapt their informatics pipelines toward assembling and annotating the genome. Parts of this process of assembly and annotation were conducted in conjunction with those who had been involved in the mapping

7. On the nature and methodological implications of archives and oral history interviews when addressing the history of recent biomedicine, see Christine Aicardi and Miguel García-Sancho, "Towards Future Archives and Historiographies of 'Big Biology,'" *Studies in History and Philosophy of Biological and Biomedical Sciences* 55 (2016): 41–44.

8. For historical research that synthesized oral histories, archival sources and published scientific literature on these collaborative projects, see note 5 and James W. E. Lowe, "Humanising and Dehumanising Pigs in Genomic and Transplantation Research," *History and Philosophy of the Life Sciences* (Forthcoming).

of the pig genome and sequencing of individual pig genes.[9] This continuity of research groups and institutions throughout mapping and whole-genome sequencing constitutes a crucial feature when characterizing the history of pig genomics and drawing broader conclusions for the wider historiography of genomics. The SGSP produced the first reference sequence of the pig genome in 2009. A reference sequence, or reference genome, is a consensus DNA sequence that constitutes the standard for a given species or subspecies and is stored in a publicly accessible database.

Rather than starting our account with one of these concerted projects, or the high-profile actors and institutions that led them, in this paper we analyze datasets, visualizations, and statistics highlighting DNA sequencing activity. We systematically collected publication data derived from every DNA sequence submission for *Sus scrofa* held by the European Nucleotide Archive from 1990 to 2015.[10] These dates were selected to capture sequencing activity from the beginning of concerted efforts to map the pig genome, until after the publication of the first reference sequence. Using institutional co-authorship relations deriving from the first publication listed for sequence submissions, we constructed a network visualization. Our aim was to discern the various *roles* played in the network by different institutions and to ascertain the factors underpinning patterns we observed of differing collaborative relationships before and after the production of the reference genome.

We used the co-authorship network to identify different *modes* of research. Further, in this paper we took the material bases underpinning co-authorship relations more into account, for example in examining collaborations that deployed particular research *tools* or *breeds* of pig. This enabled us to use our co-authorship network as a platform to investigate how the circulation and sharing of tools and resources have shaped pig genomics research. This focus complements the other papers in this special issue in demonstrating another key dimension of the history and organization of genomics that can be explored using our mixed-methods approach.

Our analysis of the changing collaborative relations around pig DNA sequencing over time reveals a picture of the development of an international community with a well-connected core of participants and institutions,

9. Lowe, "Sequencing" (n.5).

10. Rhodri Leng, Gil Viry, Miguel García-Sancho, James Lowe, Mark Wong, and Niki Vermeulen, "The Sequences and the Sequencers: What Can a Mixed-Methods Approach Reveal about the History of Genomics?," this issue.

engaged in efforts to produce and improve maps and sequences.[11] This community has largely been comprised of institutions and researchers who investigated pig genetics with the goal of developing resources that allow breeders to improve their breeding programs. It has additionally included researchers who have investigated the genetics of the pig for other purposes, such as the development of *Sus scrofa* as an animal model of human diseases, and in particular transplantation biology and practice. These researchers have often also engaged with the more agricultural concerns of the rest of the community.

Our methods have allowed us to characterize the ways in which genomics research evolves either side of a reference genome's creation. We have done this by examining the collaborative relationships exhibited by the pig genomics community over time. Through this, we have discerned longer-term developments in the structure of the community, the nature of collaborations, and modes and targets of research. In assessing the factors shaping and responding to these developments, including the circulation of tools and materials, we have been able to identify the impact and importance of the pig reference genome. We interpret the reference genome as an inflection point in the history of genomics rather than constituting its primary outcome: it enables particular forms of research and collaborations to be pursued either newly or to a greatly increased extent.

2. THE CONCEPTUAL BASIS AND CONTRIBUTION OF THIS PAPER

In this paper, we identify a shift in the nature of publications and coauthorship patterns before and after the production of the designated reference genome for *Sus scrofa* in the second half of the 2000s, with the first full draft made publicly available in 2009. This involved a shift in relative weight (rather than complete replacement) from what we term an *intensive* mode of sequencing to an *extensive* mode. Understanding and conceptualizing this shift of emphasis is historiographically important because it shows that in some nonhuman animal species, producing a reference sequence was not always considered to be the primary goal or final outcome of genomics research. Despite significant changes in their modes of research, there were substantial continuities of the communities involved in pig genomics before, during, and after the production of the *Sus scrofa* reference sequence. Instead of considering it the

11. James W. E. Lowe, "Adjusting to Precarity: How and Why the Roslin Institute Forged a Leading Role for Itself in International Networks of Pig Genomics Research," *The British Journal for the History of Science* 54, no. 4 (2021): 507–530.

culmination of their endeavor, these communities used the sequence data to open new avenues of research in the extensive rather than the intensive mode. These intensive and extensive modes of sequencing pertain to the scope and resolution that sequencers pursue. Throughout the paper, we also refer to modes of research concerning the goals guiding that pursuit. In pig genomics, two leading modes of research involve sequencing and analyzing sequence data to inform agriculture, on the one hand, and systematics (the term we use to capture the study of evolution, diversity, and domestication) on the other. While in pig genomics the intensive and extensive modes of sequencing are generally associated with the agricultural and systematic modes of research, respectively, we use separate terms for each as this relationship may not hold for genomics research concerning other species to which this framework can also be applied.

Intensification involves increasing the resolution of data concerning the genome of a given species. It can involve the identification of particular genes and other genomic features—and variants thereof—in greater numbers, and the delineation of nucleotides along ever-larger regions of chromosomes. Typically, intensification involves the use and augmentation of reference resources such as maps, mapping tools, and sequences to represent the species—in this case *Sus scrofa*.

Extensification involves the establishment of connections between the kinds of representations generated in intensive genomics and ones characterizing distinct pig breeds, families, and populations. Extensification is therefore analogous to comparative genomics, but rather than operating between abstracted representations of the genomes of less-related species (e.g., between pig and human), it instead operates between representations of closely related organisms, be they different breeds (or populations) of *Sus scrofa*, members of the Sus genus, Suidae family, or Suina suborder.[12] Extensification therefore represents the ramification of reference genomics: the reference genome produced as the culmination of intensification becomes used to help generate more specific resources with a narrower representational scope, enabling finer-grained comparisons and inferences to be made. In doing so, it has enabled more ambitious systematic surveys of diversity and evolutionary relationships.[13]

12. For a characterization of comparative genomics, see Lowe, "Humanising" (n.8).

13. The intensive mode can include horizontal and vertical sequencing, concepts we use elsewhere in this special issue to refer to the dimensionality of intended research using sequence data. Extensive sequencing may make use of the products of horizontal as well as vertical sequencing. See Miguel García-Sancho, Rhodri Leng, Gil Viry, Mark Wong, Niki Vermeulen, and James Lowe, "The Human Genome Project as a Singular Episode in the History of Genomics," this issue.

To make sense of the nature of intensification in pig genomics, and the transition to a more extensive mode of genomics, we consider the role of *bricolage*, the repurposing of resources and tools developed for other purposes. Bricolage is tinkering, and a *bricoleur* is a tinkerer. Those who tinker make do with the tools and materials available to them.[14] In this regard, the products of bricolage become inextricably connected to the activity of producing them—*bricolage* can thus operate simultaneously as a verb and a noun, designating both the process and its outcomes.

The tools and materials at the tinkerer's disposal are not custom-made for the purposes to which they may be put. They survive from previous projects. The bricoleur must combinatorially reassign what is already at their disposal in an environment where resources are limited. Bricolage is nonteleological, operating with constrained but open-ended creativity and a multiplicity of potential outcomes. There is thus historicity to its manifestations and products.[15] This historicity, along with the limited financial resources and political constraints with which pig geneticists operated—compared to other communities addressed in this special issue—make bricolage an especially suitable lens to analyze the production and use of the *Sus scrofa* genome.

The suitability of the concept of bricolage, despite it designating a rather general practice of repurposing scientific and technical resources, will become evident across the timeframe of this study. There are continuities of persons, institutions, biomaterials, and tools, along with dramatic changes in the affordances of genomics and shifts in the mode of conducting genomics. The resource limitations of the community of pig bricoleurs entail a bricolage of

14. The concept has its origins in the anthropological thought of Claude Lévi-Strauss, but our characterization is closer to the biologist François Jacob's: Claude Lévi-Strauss, *The Savage Mind* (Chicago: University of Chicago Press, 1966); François Jacob, "Evolution and Tinkering," *Science* 196, no. 4295 (1977): 1161–66.

15. Christopher Johnson, "Bricoleur and Bricolage: From Metaphor to Universal Concept," *Paragraph* 35 (2012): 355–72. Bricolage has been used in the social sciences in analyzing processes of improvisation and innovation, sometimes under conditions of scarcity: Joe L. Kincheloe, "Introduction: The Power of the Bricolage: Expanding Research Methods," in *Rigour and Complexity in Educational Research: Conceptualizing the Bricolage*, eds. Joe L. Kincheloe and Kathleen S. Berry (Maidenhead, UK: Open University Press, 2005), 1–22; François Lambotte and Dominique Meunier, "From Bricolage to Thickness: Making the Most of the Messiness of Research Narratives," *Qualitative Research in Organizations and Management: An International Journal* 8 (2013): 85–100.

existing resources that were conceived and configured—and often reconceived and reconfigured—for other purposes.[16] These include maps of various kinds, DNA libraries, mapping tools, and sequence data. This bricolage was a feature of the intensification stage and was reflected in the continuity of personnel and institutions from the mapping programs of the 1990s to the whole-genome sequencing of the 2000s. This was a distinctive feature of pig genomics, when compared with human genomics, in which the prior mapping communities comprised of medical geneticists were largely sidelined from the whole-genome sequencing project. The bricolage of pig genomics also enabled many of the same actors to operate and thrive in the subsequent extensification phase, using different tools and involving different kinds of collaborative relationships and ways of working.

These processes of bricolage and their importance suggest that rather than being based on a long-term premeditated strategy, the reference genome of *Sus scrofa* resulted from a continuous, creative, and open-ended repurposing of tools and resources necessitated by the relative lack of resources and political-institutional support of the communities involved. In this sense, the history of the production of the pig reference genome differs substantially from those of yeast and human, both addressed in the previous papers of this special issue

16. In this respect, the material practices of the pig genomics community are analogous to the "thrifty science" characteristic of much pre–nineteenth century science, treating materials as open-ended and capable of re-use and adaptation in different contexts, albeit without the moral dimension and domestic spaces associated with thriftiness; Simon Werrett, *Thrifty Science: Making the Most of Materials in the History of Experiment* (Chicago: University of Chicago Press, 2019). The material practices of bricolage in pig genomics links the history of genomics with wider themes concerning the materiality of scientific objects and practices, and the circulation of materials and associated practices, e.g., Soraya de Chadarevian, "Mapping the Worm's Genome: Tools, Networks, Patronage," in *From Molecular Genetics to Genomics: The Mapping Cultures of Twentieth-Century Genetics*, eds. Jean-Paul Gaudillière and Hans-Jörg Rheinberger (Abingdon: Routledge, 2004), 95–110; Stephen Hilgartner, "Mapping Systems and Moral Order: Constituting Property in Genome Laboratories," in *States of Knowledge: The Co-production of Science and Social Order*, ed. Sheila Jasanoff (New York: Routledge, 2004), 131–41; Robert E. Kohler, *Lords of the Fly: Drosophila Genetics and the Experimental Life* (Chicago: University of Chicago Press, 1994), chap. 5; Hanna Landecker, *Culturing Life: How Cells Became Technologies* (Cambridge, MA: Harvard University Press, 2009); Karen Rader, *Making Mice: Standardizing Animals for American Biomedical Research, 1900–1955* (Princeton, NJ: Princeton University Press, 2004). On the circulation of tools and materials beyond twentieth-century genetics and biology, see Ursula Klein and E. C. Spary, eds., *Materials and Expertise in Early Modern Europe: Between Market and Laboratory* (Chicago: University of Chicago Press, 2010); James A. Secord, "Knowledge in Transit," *Isis* 95, no. 4 (2004): 654–72.

and whose reference genome efforts received substantial political and financial backing.[17] The pig reference sequence was produced using four genome libraries containing DNA from five breeds and using resources such as radiation hybrid panels and comparative maps available from prior projects (e.g., PiGMaP and USDA-sponsored ones). Only a few of these were precision-engineered for the purpose of producing a sequence encompassing the whole *Sus scrofa* genome.

Our concept of bricolage stemmed from a visual examination of the network generated with our datasets followed by an investigation of the patterns we observed through quantitative network metrics and qualitative analysis of individual institutions. This allowed us to identify three kinds of institution, associated with distinct modes of research and patterns of collaboration: one mainly involved in agricultural research; one conducting systematic inquiries studying evolution, domestication, variation, and diversity; and hybrid institutions, which managed to be predominant in both spheres and throughout the intensive and extensive phases. The paper analyzes these types of institutions—focusing on exemplars of each mode of research—and concludes that they represent different ways of mobilizing genomic data and resources with distinct translational goals and organizational configurations.

We show how, following the production of the reference genome, the weight of work represented in our network shifts from the agriculturally inclined and hybrid institutions toward the systematic ones. As these different ways of conducting sequencing condition disparate forms of collaboration, we can account for this shift by demonstrating how the publication patterns changed before and after the advent of the reference genome, reflecting the transformation in the dominant strategies of production and use—and bricolage—of genomic resources.

We examine these processes through key institutions that our network analysis presented as noteworthy. These are:

17. On the relative poverty of agricultural scientific research, especially at the outset of this period, see Lowe, "Adjusting" (n.11); Dmitriy Myelnikov, "Cuts and the Cutting edge: British Science Funding and the Making of Animal Biotechnology in 1980s Edinburgh," *British Journal for the History of Science* 50 (2017): 701–28. There are parallels here with the contention that scientists more generally construct "doable problems," which involves an assessment, creation, and articulation of available materials, organizational and logistical capacity, and external support, such as funding. See Adele E. Clark and Joan H. Fujimura, "What Tools? Which Jobs? Why Right?," in *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, eds. Adele E. Clark and Joan H. Fujimura (Princeton, NJ: Princeton University Press, 1992), 3–44.

- two centers of the Institut National de La Recherche Agronomique (INRA) in France, one of which (INRA Castanet-Tolosan) co-produced a radiation hybrid panel for mapping genetic markers and the other (which we label CEA-INRA Jouy-en-Josas) a genome library;
- the Roslin Institute in Scotland, which collaborated inter-institutionally across agricultural and systematic research;
- Wageningen University in the Netherlands, which mixed agriculturally inclined research with systematic work on genetically distinct breeds of pig; and
- Universitat Autònoma de Barcelona (UAB), which stood out as a signal representative of the shift toward systematic research, transforming the nature of their collaborations due to their expertise and the materiality and ancestry of the pigs at their disposal.

These examples demonstrate the importance of the material, institutional, and intellectual resources and affordances produced and manifested by these institutions, and in so doing show how the process of bricolage manifested diachronically as well as synchronically. This, in turn, allows us to detail the particular features of the transition from intensive to extensive sequencing and its consequences in the practice, organization, and historicity of genomics research. Despite this shift of emphasis from intensive to extensive sequencing being also a feature of human and yeast genomics, pig genomics is a particularly appropriate object for investigating it, as it provides a clear demonstration of the effects of this transition in the network visualizations and metrics, while also featuring a continuity of institutions across both phases (including some such as UAB that exemplify the shift in a striking and informative fashion).

We proceed next by detailing the analyses we conducted on the institutional co-authorship network, which resulted in the identification of the three sets of institutions associated with particular forms of work: agricultural, systematic, and hybrid. We then devote the following sections to exploring the forms of bricolage characteristic of these. In section 4, we examine the resource and tool production of the two INRA institutes introduced above and show how these practices contributed to the highly connected nature of these institutions. Despite originally being produced for specific purposes, and not with the eventual elucidation of the reference genome in mind, these resources and tools would be adapted to contribute toward that end.

Then, in section 5, we inspect the co-authorship relationships and publications of hybrid institutions, choosing the Roslin Institute and Wageningen University as exemplars because of their notable network positions and the ways that their collaborative research practices epitomize bricolage across agricultural and systematic modes of research. The genomic capabilities these institutions had established helped them to conduct such collaborations that allowed them to bridge clusters of institutions strongly associated with distinct modes of research and sequencing. In section 6, we use UAB as a lens with which to examine the consequences of the advent of the reference genome, looking at how the nature of this institution's published research changed before and after, and how this affected its patterns of collaboration. We conclude with a reflection on how our articulation of bricolage allows us to think anew about the temporality of the history of genomics.

3. NETWORK ANALYSIS

Rather than starting with a specific set of institutions associated with a particular region or project, we wanted a global picture of institutions and their connections. We therefore began by visually examining the network of institutional co-authorship for the whole time period (1990–2015), to identify possible patterns and to try to discern what factors and processes underpinned these. The first visualization (see figure 1) is of the main component, the largest connected subnetwork in which authoring institutions are represented by circles (nodes) and co-authorship relationships by lines (known as edges) between nodes.[18]

The red and dark-blue institutions from the United States and Germany, respectively, appear spread out over the network space. By contrast, the light-blue Chinese, dark-brown South Korean, and green Japanese nodes are well concentrated and highly-interconnected within individual countries. The greater presence of these Asian institutions is a notable feature of the pig network, contrasted with the human and yeast networks, and mainly reflects the emergence of China in genomics research from the early to mid-2000s

18. Subnetworks not connected to this main component do not exceed six institutions in size and are therefore relatively insignificant, compared with the 1021 institutions present in the main component.

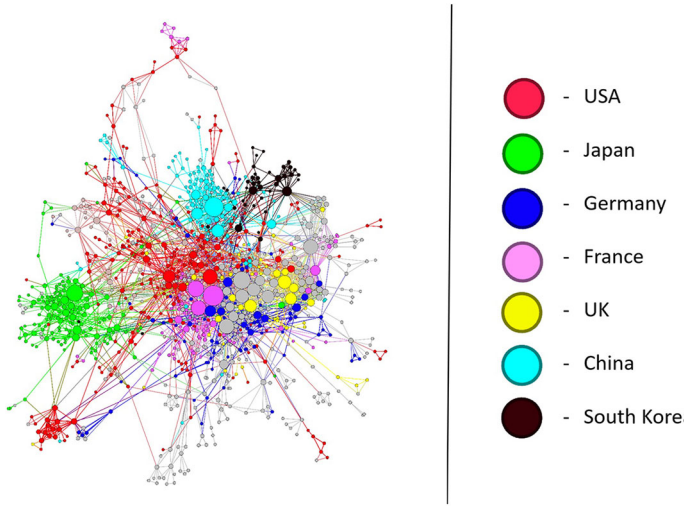

FIGURE 1. Main component of the institutional co-authorship network. We have sized the nodes to correspond to the number of publications co-authored with other institutions (weighted degree) and colored the nodes according to the country in which the represented institution is based, as indicated in the legend on the right side of the figure; we colored the rest of the nodes gray. Figure elaborated by the authors.

onward. The pig network includes publications describing sequences submitted up to 2015, while the yeast and human networks do so only for sequences submitted up to 2000 and 2005, respectively.

To further explore these apparent patterns, we generated different measures of network *centrality* using the network analysis software Gephi, which enabled us to drill down to the level of individual institutions exhibiting particular features of interest. Centrality is a measure of the connectedness of institutions to other institutions and the rest of the network. The figure for an individual institution therefore depends on its position and ties relative to the rest of the network.

We will first examine *degree centrality*, the total number of other institutions that a given institution has co-authored a publication with. Table 1 shows the top twenty-two institutions for degree centrality (due to three institutions being tied for twentieth).

TABLE 1. Top Twenty-Two Institutions for Degree Centrality*

Institution	*Country*	*Number of papers*	*Degree centrality*
Institut National de la Recherche Agronomique Castanet-Tolosan	France	52	63
Wageningen University and Research Centre	Netherlands	14	51
Kobenhavns Universitet	Denmark	38	50
Sveriges lantbruksuniversitet (Swedish Agricultural University)	Sweden	31	50
China Agricultural University	China	41	49
Institut National de la Recherche Agronomique Jouy-en-Josas	France	30	49
Universitat Autònoma de Barcelona	Spain	30	49
University of Durham	UK	6	49
National Institute of Agrobiological Sciences	Japan	37	48
Iowa State University	US	47	46
Roslin Institute	UK	16	44
CEA-INRA Jouy-en-Josas	France	33	43
Huazhong Agricultural University	China	132	43
University of Oxford	UK	7	43
USDA ARS Meat Animal Research Center	US	53	41
Museum National d'Histoire Naturelle	France	3	40
University of Aberdeen	UK	4	35
Alma Mater Studiorum Università di Bologna	Italy	16	34
Uppsala Universitet	Sweden	16	32
Ludwig-Maximilians-Universität München	Germany	13	31
Trinity College Dublin	Ireland	2	31
Universiteit Gent	Belgium	29	31

*The table shows the number of institutions that a stated institution co-authored with in the main component (degree centrality), and the number of papers underpinning these ties in our corpus. We have color-coded the institutions depending on the relationship between degree centrality and number of papers that they manifest: green indicates high numbers for both, blue indicates a high degree centrality relative to number of papers, yellow an intermediate position, whereas gray indicates an exception.

In this heavily European list, we have included the number of papers from which these ties were drawn. This sorts the top twenty-two into three sets of institutions:

1 Those, such as INRA Castanet-Tolosan (near Toulouse), that have a sizeable corpus of publications with a set of institutional co-authors nearly or at that number of papers (colored in green in table 1).

2 Those, such as Trinity College Dublin, that have a relatively high number of institutional co-authors for a smaller number of papers (colored in blue).
3 Institutions in the middle, with similar numbers of co-authors to the first group but based on a corpus of between ten and twenty papers, such as the Roslin Institute (colored in yellow). Huazhong Agricultural University in China, with by far the lowest ratio, is an anomaly (colored in gray).[19]

These different sets instantiate alternative modes of co-authorship. Set 1 institutions are characterized by co-authorship with authors from a consistent set of other institutions, reflecting longer-term collaborations and projects. Set 2 institutions appear more promiscuous, but this reflects the small number of papers associated with them in the dataset, with at least some of them having a large number of contributing authors from different institutions.

Examination of the publications authored by scientists at the top twenty-two institutions indicates that their distinct degree centralities and network properties reflect different modes of research. Set 1 institutions, such as INRA Castanet-Tolosan, Kobenhavns Universitet (University of Copenhagen), and Iowa State University, are characterized by a concentration from the early 1990s on gene mapping, mainly for the purposes of developing resources that could be used by the livestock breeding industry. We will therefore refer to this set of institutions as *agricultural*.[20]

The papers of set 2 institutions feature investigations of evolutionary relatedness and patterns of domestication and spread of different pig breeds. This work, which became prominent in the 2010s, concentrates on genetic diversity and divergence, and often requires multisite intercontinental studies, with participation often conditioned by the ability to obtain genomic resources such as the DNA of particular local breeds. The multisite nature of these studies indicates why there is such a surfeit of co-authorship links. Institutions

19. Huazhong Agricultural University is an elite institution with a large number of researchers. Many of their papers have multiple authors based there, which makes this institution less relevant in terms of co-authorship despite publishing prolifically. Authors associated with Huazhong and Chinese institutions in our dataset also tend to exhibit multiple affiliations, for instance with the Chinese Academy of Agricultural Sciences.

20. The industrial actors that translated the outputs reported in the publications are largely absent from the network, since this type of cooperation tended to materialize in internal proprietary breeding programs (and, occasionally, patent applications) rather than co-authored articles. This network absence is addressed using qualitative historical and social science methods in Ann Bruce and James W. E. Lowe, "Pigs and Chips: The Making of a Biotechnology Innovation Ecosystem," *Science and Technology Studies* (under review).

exclusively involved in these studies in the top twenty-two above are University of Durham, University of Oxford, Museum National d'Histoire Naturelle (Paris), University of Aberdeen, and Trinity College Dublin. We call these kinds of institutions *systematic*, after the biological field of systematics, which seeks to describe, classify, and explain biological diversity.

Set 3 institutions include Wageningen University and Research Centre, Alma Mater Studiorum Università di Bologna, Uppsala Universitet, Ludwig-Maximilians-Universität München, and the Roslin Institute. These co-author with the systematic institutions, as well as belonging to (and in the case of Roslin and Wageningen, leading) wider collaborations featuring the agricultural institutions. We therefore refer to these institutions as *hybrid*.

This initial quantitative analysis of network-derived data highlights two kinds of research: agricultural and systematic. There are also hybrid institutions with strong records in both systematic and agricultural research. These overlaps in modes of research reflect the fact that many leading geneticists working primarily on agriculturally motivated projects also had an interest in systematics. Work funded for primarily agricultural purposes, such as the production of resources and tools for generating and using genomic data, became of considerable use in other biological fields—such as systematic research—through processes of bricolage.

The categorization of modes of research that we propose emerged from an examination of the data on publishing patterns in conversation with the reading of the papers underlying the network ties and our prior qualitative knowledge of the co-authoring institutions and their activity. In what follows, we zoom into the network to identify case studies that exemplify those modes of research, their interactions, and evolution throughout the timeframe of our dataset (1990–2015). We examine the work of first agricultural, then hybrid, and finally systematic institutions by combining visual, qualitative, and quantitative analyses of our network, publication data, and other historical evidence. This enables us to draw conclusions about the role of these modes of research in the history of pig genomics and the historiography of genomics more generally.

4. AGRICULTURAL INSTITUTIONS: BRICOLEURS OF INTENSIVE GENOMICS

4.1. Entanglements at the Core of the Network

All the institutions we have identified as agricultural (green in table 1) were involved in large-scale pig genome mapping and sequencing projects, with the

exceptions of China Agricultural University and UAB. Based on this, and the classification into three sets, we would expect that these institutions were investigating pig genetics and genomics with the primary aim of aiding pig breeding practices, and producing tools and resources for the further exploration of the genome to that end. Moreover, and given that the successive genome projects lasted from the early 1990s to 2012, we expect that these institutions exhibit the strongest ties, compared to all of the other institutions in the network.

To examine this, we selected a subnetwork preserving only the ties underpinned by at least four co-authored publications. This figure was arrived at through testing other possible filters; four produces a sizeable enough subnetwork to make analysis worthwhile without being intractably large for a fine-grained examination.[21] This examination included a study of the kinds of research conducted, reading the published papers that form the basis of the subnetwork, and especially focusing our inquiry on the production of tools for mapping and analyzing genes within the pig genome. The qualitative analysis of the publications, as well as conducting oral histories with their co-authors and investigating their personal archives, enabled us to discern the nature of the underlying inter-institutional collaborations—in particular, what materials and tools for genome analysis they were sharing, exchanging, and circulating.

This specific focus on the material and technical sharing and exchange underpinning the co-authorship relationships offers a complementary angle to the interpretation of the other networks addressed in this special issue and highlights another benefit of our mixed-methods approach. The sharing and repurposing of pig DNA, mapping and sequencing technologies, and data is a crucial factor for explaining network positions. The prevalence of these practices also demonstrate the importance of processes of bricolage in the history of *Sus scrofa* genomics and, as we will argue, in the historiography of genomics research more generally.

After filtering to remove ties based on three co-authored publications or fewer, one sizeable subnetwork remains, the other institutions becoming isolated (see figure 2).

21. Four co-authored publications over 25 years is a mean of one co-authored publication every 6.25 years. In reality, the publications were not evenly spaced out.

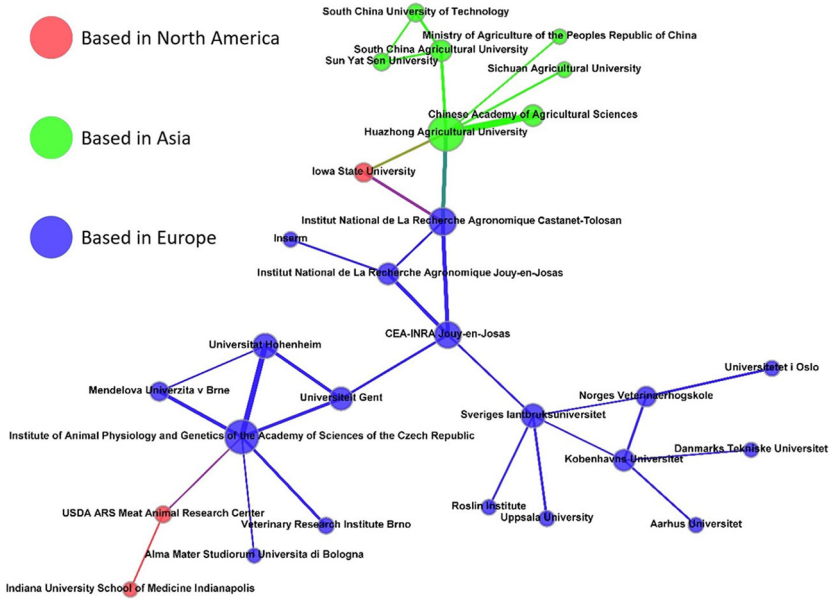

FIGURE 2. Subnetwork of the main component. We filtered it to maintain only those ties underpinned by at least four co-authored publications concerning pig DNA sequences. Figure elaborated by the authors.

The subnetwork has a French core. Attached to it are three arms, characterized by ties with particular geographical flavors: a Scandinavian/northern European one, a Chinese one, and a central European one. Three US institutions are associated with the central European arm, or between the French core and the Chinese arm. The strongest co-authorship ties show a clear geographic pattern: of intra-country co-authorship in China, co-authorship with more geographically proximate institutions within Europe, and the less geographically constrained co-authorships of US institutions, tallying with the apparent patterns in figure 1.[22]

Of the eleven institutions identified as *agricultural* in section 2, all but three (China Agricultural University, UAB, and National Institute of Agrobiological Sciences) are present in the filtered subnetwork; of these three absent

22. As shown elsewhere in this special issue, when we compared the main components of the human, yeast, and pig co-authorship networks, the latter exhibited a distinct pattern of national and transnational connections; Leng et al., "The Sequences" (n.10). We document below how the resource constraints of some areas of pig genetics research have encouraged collaboration and resource sharing, thus contributing to this distinct pattern.

institutions, the first two were not involved in international genome projects. Three of the five hybrid institutions are also present (Roslin, Bologna, and Uppsala). Many of the nodes in the subnetwork represent institutions with a strong agricultural orientation in their pig research. Indeed, the core of the co-authorship visualization is dominated by agriculturally inclined institutions, most of which were involved in concerted genomics initiatives. Thirteen of the twenty-eight institutions in the subnetwork were involved in PiGMaP, the USDA-sponsored projects or the Swine Genome Sequencing Consortium (all European except for Iowa State University and the USDA Meat Animal Research Center), displaying a remarkable concordance between publication prominence—as determined by membership of this subnetwork—and the playing of significant roles in genome projects that is not the case for the analogous human network and only partially so for the yeast network.[23]

To explore the kinds of research involved in this subnetwork further, we moved beyond the visual analysis to examine the 133 publications that underpin it. We extracted the PubMed IDs (PMIDs) from our dataset, allowing us to search for the papers using the PubMed database. Analyzing the distribution of the papers across the years, we observed that there were few papers in the 1990s, a rise in the early 2000s, and a peak in the mid-2000s, followed by a drop-off. This mirrored the pattern for all papers in the pig dataset and meant our sample was mainly formed of publications that appeared before the release of the *Sus scrofa* reference genome.

A qualitative examination of the papers for our filtered subnetwork showed that a high proportion were concerned with mapping, localization, and characterization of pig genes and their variants.[24] Broadly, these genes were related to livestock production traits such as reproductive prolificacy and fat deposition, but encompassed a wider range of functions, including immune response and basic physiological and developmental processes. Notable exponents of this work in the filtered subnetwork include Max Rothschild from Iowa State

23. García-Sancho et al., "The Human Genome Project" (n.13); Miguel García-Sancho, James Lowe, Gil Viry, Rhodri Leng, Mark Wong, and Niki Vermeulen, "Yeast Sequencing: 'Network' Genomics and Institutional Bridges," this issue. The list of participants in PiGMap is from Alan Archibald's personal papers; Partition—"EC PiGMaPII—Final Report," obtained 15 May 2017.

24. This does not necessarily mean that gene hunting was the only or main activity of these institutions, but that this was an activity that involved DNA sequence deposition and often led to co-authored publications. For a more detailed characterization of the research involving these institutions—centered on genetic markers rather than genes themselves—see Lowe and Bruce, "Genetics" (n.5).

University, who contributed to identifying candidate genes intended to be taken up by private breeding companies and used in the design of their breeding programs.[25] Rothschild was involved in the wider pig genome mapping effort (PiGMaP) as an external collaborator of the European consortium. He was also the USDA's National Swine Genome Coordinator (1993–2013) and played a leading role in helping to develop resources such as primers and probes, which he distributed to pig geneticists across the world for use in their research.

The journals in which these papers were published reflect their agricultural focus and the communities with which their authors were aiming to speak. Overall, 46 of the 133 papers underpinning the subnetwork appeared in *Animal Genetics*, a journal once edited by PiGMaP coordinator Alan Archibald of the Roslin Institute, and concerned primarily with livestock species. *Mammalian Genome*, the second most popular journal in this subnetwork, is more associated with mouse and human genetics, betraying its origins as a mouse genetics journal. Appropriately, therefore, this journal was home to many of the papers using comparative approaches and resources in gene mapping—such as using primers and cDNA from well-characterized genes, mainly from humans, as tools to pinpoint specific locations in the pig genome (e.g., in our dataset: PMIDs 15389320 and 16176575).

The peak in number of publications occurred at the outset of the project to fully sequence the genome of the pig in 2006. Subsequently, there was a decline. How can this be explained? The quantity of available sequence data, annotations of it, and data on polymorphic markers across the genome (including microsatellites and single-nucleotide polymorphisms) was rapidly increasing from the mid-2000s onward. The results were included in publicly available (or in the case of a Sino-Danish collaboration, privately accessible) databases, helping to raise the bar for what constituted a publishable contribution. This meant that intensive sequencing needed to be augmented in some way to constitute a significant research contribution. This could be achieved by studying a range of polymorphisms and genes in wider genomic regions. Alternatively, more functional research could be pursued, for example

25. E.g., in our dataset: PMID 7985839. For an overview of collaborations with industry that are generally absent from the network, see Max F. Rothschild and Graham S. Plastow, "Development of a Genetic Marker for Litter Size in the Pig: a Case Study," in *Intellectual Property Rights in Animal Breeding and Genetics*, ed. Max F. Rothschild and Scott Newman (CABI Publishing, 2002), 179–96. On the relation of publicly funded institutions involved in pig genomics with pig breeding companies, see Bruce and Lowe, "Pigs and Chips" (n.20).

analyzing expression patterns using the products of DNA transcription. Additionally, different breeds and populations of pigs could be examined: a more extensive (systematic) mode of work whose presence in the network, as we will show, expanded considerably after the release of the reference genome.

The existence of the reference genome therefore prompted and provided the basis for an extensive mode of sequencing. It encouraged different configurations of bricoleurs for distinct bricolages directed toward systematic research that included specific examinations of particular breeds, either as an alternative to intensive sequencing or—to a lesser extent—as a new form of it. The new configurations included institutions that were not as prominent in the intensive mode—those that could access breeds, families, and populations of pigs not sufficiently represented in the available *Sus scrofa* genome, despite in some cases having more modest data-production capacities. Some agriculturally inclined institutions that were prominent in the intensive mode of genomics—including those who had harbored long-term interests in local breeds, in mating distantly related breeds, or in exploring patterns of domestication—also entered into extensive sequencing.

These agriculturally oriented mappers were useful not only for their expertise in pig genetics and genomics but also for their experience as bricoleurs of a panoply of materials and tools that they had used, re-used, created, and adapted. We now focus on two bricoleurs, French institutions funded by the INRA that created resources and tools for their own research purposes. These products contributed to their collaborative entanglements with other institutions, thus explaining their network position and connections. Further bricolage saw these resources and tools used in the physical mapping and then sequencing of the whole genome of the pig. The sharing, adaptation, and repurposing of these resources by the wider pig genetics community helped shape the work and relationships of the institutions that produced them.

4.2. Intensive Resources: A Bacterial Artificial Chromosome Library and Radiation Hybrid Panel

CEA-INRA Jouy-en-Josas is an institution that piquantly illustrates the role of the production, distribution, circulation, use, and repurposing of materials in pig genomics in the intensive mode. Additionally, its position at the heart of this subnetwork can be explained by the provision of materials, as well as technical expertise developed as part of a unique mission. CEA-INRA Jouy-en-Josas denotes a laboratory jointly funded by the INRA and CEA—the

French Commissariat à l'énergie atomique (Atomic Energy Commission)—hence its common designation as the *Laboratoire mixte,* and nickname: "the atomic farm."[26] CEA was interested in biological and medical research from its inception in 1945. Alongside other national atomic energy organizations, it recognized the need to study the effects of radiation exposure, as well as exploring the potential of radioisotopes as molecular tracers and sources of genetic mutation that could be used in breeding or crossing experiments.[27]

The *Laboratoire mixte* was established in 1964, next to the INRA campus in Jouy-en-Josas, south of Paris. Under the leadership of Marcel Vaiman, they were tasked with investigating the immune genetics of pigs, initially because of the use of the pig as a model for transplantation biology and practice. This research was originally supported by CEA to test the effects of gamma radiation and to effect bone marrow transplants to replace irradiated tissue. In 1970, Vaiman and his team (including Christine Renard, who was to be a long-serving member of the laboratory; Patrick Chardon, another key member, joined later) demonstrated that, as well as humans and mice, the pig possesses a Major Histocompatibility Complex (the Swine Leucocyte Antigen system; SLA), a set of densely packed and highly variable genes involved in immune response.[28]

Oral histories and archival research we conducted in Paris and at Jouy-en-Josas revealed three linked roles that account for the team's central position in

26. Claire Rogel-Gaillard, personal communication, December 2021. While we refer to it as CEA-INRA Jouy-en-Josas or *Laboratoire mixte* for ease of reference in this paper, it mainly went by two names across its history: Laboratoire de Radiobiologie Appliquée (Laboratory of Applied Radiobiology) and Laboratoire de Radiobiologie et Etude du Génome (Laboratory of Radiobiology and Genome Study).

27. Matthew Adamson, "Cores of Production: Reactors and Radioisotopes in France," *Dynamis* 29 (2009): 261–84; Angela N. H. Creager, *Life Atomic: A History of Radioisotopes in Science and Medicine* (Chicago: University of Chicago Press, 2013). These research agendas are inseparable from the wider attempt to portray the outcomes of atomic research as peaceful and constructive: John Krige, "Atoms for Peace, Scientific Internationalism, and Scientific Intelligence," *Osiris* 21, no. 1 (2006): 161–81; Nicolas Rasmussen, "The Mid-century Biophysics Bubble: Hiroshima and the Biological Revolution in America, Revisited," *History of Science* 35, no. 3 (1997): 245–93.

28. Marcel Vaiman, Christine Renard, Philippe LaFage, Jacques Ameteau, and Pierre Nizza, "Evidence for a Histocompatibility System in Swine (SL-A)," *Transplantation* 10 (1970): 155–64. Much of this pioneering research was carried out in the midst of interactions with Jean Dausset, one of the discoverers of the Human Leucocyte Antigen system, and his team. On the Human Leucocyte Antigen system and its place in the history of genetics, see Peter S. Harper, *A Short History of Medical Genetics* (Oxford: Oxford University Press, 2008), 201–8. For a more extensive discussion of CEA-INRA Jouy-en-Josas in the contexts of transplantation biology and pig genomics, see, respectively, Lowe, "Humanising" (n.8) and Miguel García-Sancho and James W.E. Lowe, *A History of Genomics Across Species, Communities and Projects* (Palgrave Macmillan, forthcoming), chap. 5.

the subnetwork: resource creation and development, fundamental research, and a service function. As well as conducting research into the immunogenetics of the pig, Vaiman and his colleagues provided serological typing services based on this research and Chardon led the creation of a Bacterial Artificial Chromosome (BAC) library. Claire Rogel-Gaillard joined the group to work on constructing genome libraries. A genome library or a DNA library is a material resource in which cloned fragments of DNA are stored in a host vector, such as *E. coli* bacteria in the case of BACs. The motive for constructing the BAC library was to map genes (especially those in the SLA) and to try to identify whether functional retroviral sequences—that might be detrimental to the potential use of pigs as organ donors—were present in the genome. Using this library, in the early 2000s Vaiman's team conducted the first major sequencing study of a region of the swine genome (the SLA complex), in conjunction with the French large-scale sequencing center, Genoscope.[29] This represented a contribution toward the development of the pig as a biomedical model, one of the stated motivations of sequencing the whole genome.[30]

As managing the library involved work too arduous for merely looking at one genomic region, Vaiman's team endeavored to share this resource with colleagues across the world. Due to the logistics of storing and screening libraries for multiple species, the INRA BAC-YAC Resource Center was established to manage this and other libraries produced by the *Laboratoire mixte* and to distribute clones on request. There, full-time staff including technicians and scientists would process requests and screen the libraries for relevant clones to be sent out, initially free of charge.[31] The provision of clones to other groups sometimes sparked collaborations, for example in the localization and analysis of the genomic region associated with the Rendement Napole allele linked to poorer meat quality (in our dataset: PMID 11401445). The materials and capabilities developed at the *Laboratoire mixte*, including the services they

29. Christine Renard, Elizabeth Hart, Harminder Sehra, Helen Beasley, Penny Coggill, Kerstin Howe, Jennifer Harrow, et al., "The Genomic Sequence and Analysis of the Swine Major Histocompatibility Complex," *Genomics* 88 (2006): 99–110; interview with Patrick Chardon, Christine Renard, and Marcel Vaiman, Paris, conducted by James Lowe, 28 November 2017. Genoscope, which was also part of the CEA, is a leading submitter of both pig and human sequences in our dataset, and became a member of the International Human Genome Sequencing Consortium: García-Sancho and Lowe, *A History* (n.28), chap. 4.

30. Gary Rohrer, Jonathan E. Beever, Max F. Rothschild, Lawrence Schook, Richard Gibbs, and George Weinstock, "Porcine Sequencing White Paper: Porcine Genomic Sequencing Initiative" (2002). https://core.ac.uk/download/pdf/212813224.pdf

31. Interview with Claire Rogel-Gaillard, conducted by James Lowe over Skype, 3 May 2017.

provided, the data they produced, and the clones they distributed, all account for this remarkable laboratory's position at the heart of the network, and explain how they bridge so many different arms of it.

The BAC library that Vaiman's team had compiled for immunological research and the sequencing of the SLA region was used in the construction of a physical map of the pig genome as a prelude to the whole-genome sequencing effort. In turn, the physical map could be used by researchers to identify which clones they would like to request from the Resource Center. The production of this library is a salient example of how expertise and resources developed for one form of research (immunology) became useful in larger-scale genomic projects through bricolage. Pig genomics involved more bricolage than human genomics, for which all of the libraries and maps used in sequencing were produced explicitly for the purpose, thanks to the privileged political and financial support of the international consortium in charge of the production of the reference genome of *Homo sapiens*. Yeast genomics was closer to pig genomics than human genomics, with some materials (such as DNA libraries and strains) being repurposed and others produced anew for dedicated, whole-genome sequencing projects.[32]

For the physical mapping of the pig genome, four BAC libraries were used in conjunction with another key resource: a radiation hybrid panel called IMpRH (INRA-Minnesota porcine Radiation Hybrid) developed from the late 1990s, initially through a collaboration between another INRA station at Castanet-Tolosan and the University of Minnesota. This reflected a long-standing partnership between INRA Castanet-Tolosan and Lawrence Schook, who in 2000 moved from Minnesota to the University of Illinois, Urbana-Champaign, and later became the joint-head of the Swine Genome Sequencing Consortium that led the SGSP.[33]

32. On the International Human Genome Mapping Consortium and its role in the production of the reference sequence of *Homo sapiens,* see García-Sancho and Lowe, *A History* (n.28), chap. 4. On yeast genomics, see García-Sancho et al., "Yeast Sequencing" (n.23).

33. Martine Yerle, Philippe Pinton, Annie Robic, A. Alfonso, Y. Palvadeau, Chantal Delcros, Rachel Hawken, et al., "Construction of a Whole-Genome Radiation Hybrid Panel for High-Resolution Gene Mapping in Pigs," *Cytogenetics and Cell Genetics* 82 (1998): 182–88. This move by Schook's group probably explains the absence of the University of Minnesota and the University of Illinois, Urbana-Champaign, from the filtered subnetwork. Only a single paper links the University of Minnesota with INRA Castanet-Tolosan (in our dataset: PMID 14970687). This indicates, though, that a tie underpinned by a single paper sometimes enables us to see shadows of significant collaborations that may then be further investigated.

A radiation hybrid panel is a mapping tool based on the use of radiation to break apart DNA into fragments. The frequency of the co-incidence of markers on these fragments can then be measured and analyzed to map the relative positions of the markers and the distances between them. IMpRH was used to produce a comparative map between the human and pig genomes,[34] which also aided subsequent genomic research, including the production of a physical map and identification of clones for sequencing as part of the SGSP.[35] IMpRH was accompanied by a web server on which users could access previously entered data and input mapping information on their own markers. By 2006, the year in which the SGSP formally started, it had submitter and submission data for 7,138 markers, alongside other information.[36]

In total, 49 of the 133 papers underpinning the subnetwork used the IMpRH in some way. Our analysis of these papers shows that the use of the tool was especially vital to the work of the Institute of Animal Physiology and Genetics of the Academy of Sciences of the Czech Republic, CEA-INRA, and Huazhong Agricultural University. Indeed, all of the papers linking INRA Castanet-Tolosan with Huazhong Agricultural University describe work that used IMpRH. Martine Yerle, a key researcher involved in the creation of IMpRH, was the French co-author on all of these papers. Several materials and methods sections indicated the use of Chinese breeds, such as Tongcheng pigs. Co-authors were therefore brought together by tools (IMpRH) and material contributions of animals with distinctive characteristics, genetic or otherwise. Data, knowledge, and tools circulated from one institution to another, and this process materialized in the co-authored publications.

The bricolage performed in agricultural institutions shows that the use of certain resources and tools may not be the same as the one envisaged when they were produced. IMpRH was a tool originally conceived for the concrete goal of mapping genes and genetic markers for selective breeding. The transformation of this tool into a genomic resource was thus more open-ended and emergent

34. Stacey N. Meyers, Margarita B. Rogatcheva, Denis M. Larkin, Martine Yerle, Denis Milan, Rachel J. Hawken, Lawrence B. Schook, and Jonathan E. Beever, "Piggy-BACing the Human Genome II: A High-Resolution, Physically Anchored, Comparative Map of the Porcine Autosomes," *Genomics* 86 (2005): 739–52.

35. Lowe, "Sequencing" (n.5), 17.

36. Denis Milan, Rachel Hawken, Cédric Cabau, Sophie Leroux, Carine Genet, Yvette Lahbib, Gwenola Tosser, et al., "IMpRH Server: An RH Mapping Server Available on the Web," *Bioinformatics* 16 (2002): 558–59. This resource is currently inactive online; after being directed to it by Denis Milan while conducting our research in France, we accessed it and compiled its contents.

than post-hoc reconstructions centered on large-scale sequencing projects might allow. Appreciating this open-endedness of bricolage provides a means of challenging accounts of genomics that teleologically posit the eventual production of a reference genome as an end in itself that motivates prior activity. Instead, at every point in the process, the repurposing of resources was shaped by concrete research aims that were not necessarily viewed as mere preliminaries to the sequencing of a reference genome. As we now show, this open-endedness meant that institutions that had a general agricultural orientation in their mapping and sequencing work were also able to pursue other modes of research and explore other ways of characterizing the genomes of pigs.

5. HYBRID INSTITUTIONS: BRICOLEURS OF DIFFERENT MODES OF RESEARCH

The Roslin Institute and Wageningen University played key roles in PiGMaP and other large-scale projects. In light of this, the peripheral position of the Roslin Institute in the filtered subnetwork and the absence altogether of Wageningen University is surprising. The figures we showed for the overall number of publications and distinct co-authorships for these institutions in section 3 provide one clue toward an explanation of this. The number of distinct institutions that they co-authored with was higher relative to their number of papers than the Set 1 institutions that included many of the other agriculturally inclined institutions that participated in projects with the Roslin Institute and Wageningen University. This indicates that their collaborative activity was more wide-ranging than many of the institutions in the filtered subnetwork. After demonstrating this below by examining some of the publications that underpinned co-authorship connections of the Roslin Institute and Wageningen University, we highlight how they bridge different parts of the network. These different parts of the network are populated by institutions carrying out different forms of research, such as intensive and agriculturally oriented, or extensive and systematic.

Wageningen and Roslin are both mainly agriculturally inclined in their research. Most of their papers in the dataset involve the characterization of genes, including the identification of polymorphisms and mutations thought to be associated with variation in livestock production traits such as backfat thickness. Both Wageningen and Roslin were also bricoleurs of maps,

materials, and tools they themselves had helped to create and that were later repurposed from the characterization of agriculturally relevant genes to the determination of the reference genome. Yet the direction of Wageningen and Roslin's bricolage did not always relate to the eventual construction of the reference genome.

One of the Roslin co-authorships is a study of pig domestication using mitochondrial DNA extracted from domesticated pigs and wild boar from all around the world (in our dataset: PMID 15761152). Two of the Wageningen collaborative articles are concerned with diversity: one examined the genetic diversity of rare breed Chato Murciano pigs in Spain (in our dataset: PMID 23051150); another was a study of pig domestication and dispersal using ancient mitochondrial DNA (in our dataset: PMID 23180578). These publications required inter-institutional collaboration—the first and third with a considerable variety of institutions—because of the systematic mode of research involved; we reflect on the historiographical implications of this below.

The latter study that Wageningen was involved in, published in 2013, constructed phylogenetic trees using DNA sequences obtained from modern wild boar living in the middle-east, in addition to other modern pig DNA sequences previously submitted to GenBank.[37] The sequencing work, therefore, depended upon prior sequence submissions so that robust conclusions could be drawn from comparative analyses using newly generated sequences, and on collaboration with institutions in multiple continents. Among those preexisting sequence records, the reference genome, which had been published in the literature one year earlier, in 2012, played a pivotal role: it served as the fundamental scaffold on which new sequence data could be put together, and to which it could refer. This points to a bricolage of the reference sequence itself toward systematist goals and a shift of emphasis from intensive to extensive sequencing: from compiling a reference genome sequence to using it in sequence comparisons across specific pig breeds, families, and populations.

The visualizations below demonstrate that the metrics that placed Wageningen (see figure 3) and Roslin (see figure 4) in the hybrid set are further indications of the particular roles that these institutions have played in the full co-authorship network. For both of them, their co-authorship ties link

37. For example, in studies that themselves used previously submitted sequences in their analyses, such as Greger Larson, Umberto Albarella, Keith Dobney, Peter Rowley-Conwy, Jörg Schibler, Anne Tresset, Jean-Denis Vigne, et al., "Ancient DNA, Pig Domestication, and the Spread of the Neolithic into Europe," *Proceedings of the National Academy of Sciences* 104 (2007): 15276–81.

a particular arm of the filtered subnetwork (agricultural nodes, on the right side of figure 3, left of figure 4) to those we identified as members of the systematic set (on the left side of figure 3, right of figure 4). In both cases, although they have minimal or zero ties underpinned by at least four co-authored papers, they show a wide number of connections with institutions across five continent regions in the case of Wageningen, and four for Roslin.

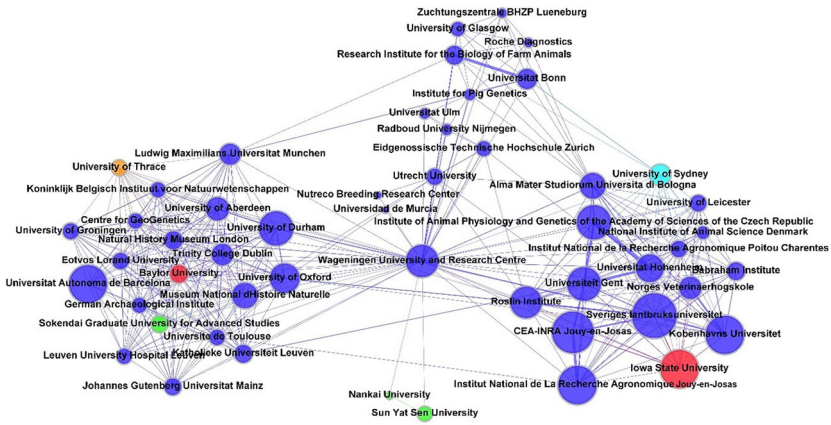

FIGURE 3. Network diagram depicting all co-authorship ties of Wageningen University. We adjusted the spatial configuration of some nodes for ease of visualization. Figure elaborated by the authors.

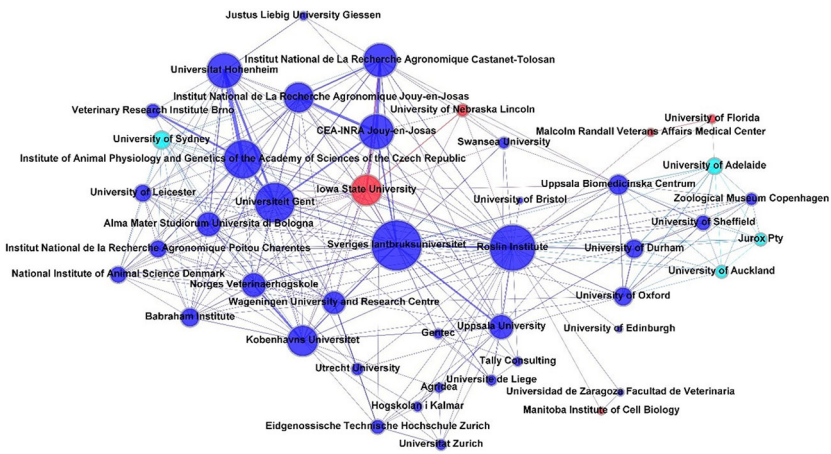

FIGURE 4. Network diagram depicting all co-authorship ties of the Roslin Institute. We adjusted the spatial configuration of some nodes for ease of visualization. Figure elaborated by the authors.

The network positions and underlying research practices of Roslin and Wageningen exemplify the role of hybrid institutions both within our co-authorship network and the history of genomics more generally: to bridge different research traditions and the sets of practices associated with them. Their absence or marginality in the filtered subnetwork reflects the temporary nature of many of the collaborations that these institutions embarked upon. As the 2010s progressed, Roslin and Wageningen broadened their focus beyond their original agricultural concentrations, becoming responsive to and seeking out systematic and also biomedically inclined partners.[38] The analytical lens of bricolage helps us to apprehend and interpret these moves, the co-authorship ties that resulted beyond the network core, and the broader political and funding conditions under which they occurred.

Collaborations bridging different modes of research necessitated the sensibility of the bricoleur, encouraging the imaginative retooling of existing data, models, materials, and methods, and at the same time shifting the balance of the bricolage of pig genomics from agricultural to systematic.[39] In contrast to *Homo sapiens*, the production of the pig reference genome did not involve the construction of a *thin* domain separate from existing domains of biological research purely focused on large-scale, whole-genome mapping and sequencing. Instead, pig reference genomics performed a *thicker* connective or bridging function. This thicker breadth and greater porosity was activated and continually produced through the sensibility and activity of the pig bricoleurs.

38. For example, in our network the Roslin Institute has co-authorship ties with the Department of Medicine at the University of Florida and the Malcolm Randall Veterans Affairs Medical Center, as well as cancer researchers at the University of Edinburgh (in our dataset: PMIDs 15718102 and 18465210). This responsive disposition was provoked by the precarious funding and public support that many agriculturally inclined institutions experienced from the 1980s onward; see Myelnikov, "Cuts" (n.17); Lowe, "Adjusting" (n.11). On the status of agriculture at land grant universities in the 1990s against the backdrop of their institutional history dating back to the 1860s, see National Research Council, *Colleges of Agriculture at the Land Grant Universities: A Profile* (Washington, DC: The National Academies Press, 1995).

39. This resonates with another example of large-scale shifts in research, Christophe Bonneuil and Frédéric Thomas's portrayal of INRA as a network of institutes shifting their priorities from the genetic improvement of plants to the use of transgenic technologies. In this changing research agenda, collaboration between different types of INRA institutes was as important as the shifting internal objectives of each one: Christophe Bonneuil and Frédéric Thomas, *Gènes, pouvoirs et profits* (Versailles, France: Éditions Quae, 2009). It is clear, however, that the increased availability of reference sequence and associated data was the crucial factor in the shift we describe here.

Hybrid institutions, in connecting agriculturally and systematically oriented research through their co-authorship ties, also bridge the histories of intensive and extensive sequencing. The creation of maps, materials, resources, and a reference genome furthered the intensive sequencing mode while creating the means for an extensive sequencing to flower as a ramification of genomics research. For these resources to be repurposed, they needed to be constructed in such a way so as to be sufficiently abstracted or separated from the motivations and circumstances of their creation. This enabled other actors, with different purposes and approaches, to use them in a way that suited their goals. However, as we have seen, repurposing often involved co-authorship with the institutions that generated the resources in the first place. Further, in the case of hybrid institutions, they exhibit the disposition to pursue new lines of research based on the opportunity presented by a laboratory approaching them with work distinct from what they have done before.

Looking at the relationships between hybrid, systematic, and biomedically inclined institutions shows the importance of not reducing the history of pig genomics to agricultural research and associated large-scale genome projects. Instead, it is vital to observe and examine the ways in which these key endeavors in the history of pig genomics are connected to other modes of research and institutions that are more peripheral in the network. The continuities of people and institutions involved in pig genomics before, during, and after the construction of the reference genome provides a good opportunity for discerning and analyzing dimensions of bricolage that may be more difficult to draw out of the yeast and human networks. The difficult political and financial conditions experienced by many agriculturally oriented research institutions from the 1980s onward engendered institutional dispositions to the diversification of research and collaborative activities. By displacing our analysis from the more central agricultural and hybrid institutions to more peripheral systematic ones, we can further our understanding of this process of diversification of research, in so doing making bricolage and the open-endedness of the history of pig genomics more evident.

6. THE SHIFT TO SYSTEMATIC RESEARCH: BRICOLAGE FOR AND IN EXTENSIVE GENOMICS

We have portrayed the pig reference genome as being the result of bricolage of tools, data, and material resources initially deployed for agricultural and

immunogenetic research purposes. We now examine whether the advent of the reference genome encouraged new forms of work by different sets of bricoleurs. One way to assess this is to examine collaborative sequencing and publishing activity before and after the first full draft version of the reference genome of the pig, Sscrofa9, which became available in November 2009. This enables us to further explore the temporal, diachronic dimension of the network and see whether it acquired different configurations across the period addressed by our dataset (1990–2015). Given that we have already used degree centrality to identify our different sets of institutions, we draw on two other centrality scores here: *closeness*—the mean distance of a given institution from all other institutions in the network—and *betweenness*, a measure of the number of times a given institution lies on the shortest path between two other institutions.

First, we present the top twenty institutions as ranked according to closeness centrality for the two periods (table 2) and, second, we list the top twenty institutions as ranked according to betweenness centrality for the two periods (table 3):

TABLE 2. Leading Closeness Centralities over Successive Periods*

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
1	INRA Castanet-Tolosan	France	Europe	China Agricultural University	China	Asia
2	CEA-INRA Jouy-en-Josas	France	Europe	Huazhong Agricultural University	China	Asia
3	Iowa State University	US	North America	University of Aberdeen	UK	Europe
4	Sveriges lantbruksuniversitet	Sweden	Europe	University of Durham	UK	Europe
5	USDA ARS Meat Animal Research Center	US	North America	Ministry of Agriculture of the Peoples Republic of China	China	Asia
6	Kobenhavns Universitet	Denmark	Europe	University of California Davis	US	North America

(continued)

TABLE 2. *(continued)*

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
7	INRA Jouy-en-Josas	France	Europe	Jiangxi Agricultural University	China	Asia
8	University of Sydney	Australia	Oceania	Universitat Autònoma de Barcelona	Spain	Europe
9	Universiteit Gent	Belgium	Europe	Chinese Academy of Agricultural Sciences	China	Asia
10	Roslin Institute	UK	Europe	Chinese Academy of Sciences	China	Asia
11	Alma Mater Studiorum Università di Bologna	Italy	Europe	Chinese Academy of Social Sciences	China	Asia
12	Universität Göttingen	Germany	Europe	Hubei Provincial Institute of Cultural Relics and Archaeology	China	Asia
13	Institute of Animal Physiology and Genetics of the Academy of Sciences of the Czech Republic	Czech Republic	Europe	Katmai National Park	US	North America
14	Wageningen University and Research Centre	Netherlands	Europe	Lanzhou University	China	Asia
15	Universität Hohenheim	Germany	Europe	University College London	UK	Europe
16	USDA ARS Beltsville Human Nutrition Research Center	US	North America	Uppsala Universitet	Sweden	Europe
17	Uppsala Universitet	Sweden	Europe		China	Asia

(continued)

TABLE 2. *(continued)*

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
				Sichuan Agricultural University		
18	National Institute of Agrobiological Sciences	Japan	Asia	University of Kentucky	US	North America
19	Norges veterinærhøgskole (Norwegian School of Veterinary Science)	Norway	Europe	South China Agricultural University	China	Asia
20	INRA Tours	France	Europe	Peking University	China	Asia

*The periods cover either side of the online release of the pig's draft reference genome in 2009. We color-coded the institutions according to whether their publications in that period were mainly agriculturally inclined (green), mainly systematic (pink), or mixed/other (blue—the category of other mainly denotes biomedical science, immunology or molecular biology).

TABLE 3. Leading Betweenness Centralities over Successive Periods

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
1	INRA Castanet-Tolosan	France	Europe	Huazhong Agricultural University	China	Asia
2	USDA ARS Meat Animal Research Center	US	North America	China Agricultural University	China	Asia
3	Iowa State University	US	North America	Chinese Academy of Sciences	China	Asia
4	CEA-INRA Jouy-en-Josas	France	Europe	Universitat Autònoma de Barcelona	Spain	Europe
5		Japan	Asia	University of Aberdeen	UK	Europe

(continued)

TABLE 3. *(continued)*

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
	National Institute of Agrobiological Sciences					
6	Sveriges lantbruksuniversitet	Sweden	Europe	University of Durham	UK	Europe
7	China Agricultural University	China	Asia	University of California Davis	US	North America
8	Kobenhavns Universitet	Denmark	Europe	Jiangxi Agricultural University	China	Asia
9	INRA Jouy-en-Josas	France	Europe	Chungbuk National University	South Korea	Asia
10	Universität Göttingen	Germany	Europe	Seoul National University	South Korea	Asia
11	Roslin Institute	UK	Europe	INRA Castanet-Tolosan	France	Europe
12	Kyoto University	Japan	Asia	Museum National d'Histoire Naturelle	France	Europe
13	USDA ARS Beltsville Human Nutrition Research Center	US	North America	Ministry of Agriculture of the Peoples Republic of China	China	Asia
14	Technische Universität München	Germany	Europe	Ludwig-Maximilians-Universität München	Germany	Europe
15	Oklahoma State University Stillwater	US	North America	Peking University	China	Asia
16	Tierärztliche Hochschule Hannover (University of Veterinary Medicine Hanover)	Germany	Europe	Chinese Academy of Agricultural Sciences	China	Asia

(continued)

TABLE 3. *(continued)*

	Period 1: 1990–2009			Period 2: 2010–2015		
Rank	**Institution**	**Country**	**Region**	**Institution**	**Country**	**Region**
17	Université de Montréal	Canada	North America	Sichuan Agricultural University	China	Asia
18	Universidad Autónoma de Madrid	Spain	Europe	University of Missouri System	US	North America
19	University of Minnesota System	US	North America	INRA Jouy-en-Josas	France	Europe
20	University of Tokyo	Japan	Asia	Rural Development Administration	South Korea	Asia

*The periods cover either side of the online release of the pig's draft reference genome in 2009. We color-coded the institutions according to whether their publications in that period were mainly agriculturally inclined (green), mainly systematic (pink), or mixed/other (blue—the category of other mainly denotes biomedical science, immunology or molecular biology).

With the caveat that the 2010–2015 period data relies on relatively much smaller numbers of publications for each individual institution,[40] several interesting features can be discerned. One is the shift in geographical center of gravity, particularly the rise of Asian institutions and the relative decline in the numbers of North American and European institutions. Within Asia, China and South Korea rose to prominence, while Japanese institutions dropped out of the top twenty. In Europe, Germany and Scandinavia declined, France and the UK increased their representation in the top twenty, and Spain maintained its single placing. There is also a manifest shift from agricultural and hybrid institutions toward the systematic, at least for the European institutions concerned. The highest ranked European institution over 2010–2015 for betweenness centrality, also in the top twenty for closeness centrality, is the UAB. Although its overall number of papers and degree centrality ratio put it in the agricultural set of institutions in section 2 (above), a qualitative analysis of the underlying publications points to a hybridity in UAB's institutional behavior much like Wageningen University's.

40. In the first batch, eighteen and seventeen institutions in the top twenty rankings for closeness and betweenness respectively have at least ten publications; for the second batch, the figures are six and seven, respectively.

UAB's papers in our dataset fall into two periods that match the shift in emphasis from intensive to extensive sequencing we have observed. The first period, from the late 1990s to the mid- to late 2000s, concerns papers published in journals associated with livestock genetics and geneticists: mostly *Animal Genetics* but also *Mammalian Genome*. The papers in question are mostly in the *Brief Notes* section of short reports in *Animal Genetics*, predominantly analyses of variants of particular genes of interest (to livestock production, typically). This involved intensive sequencing of a gene or genes in multiple pigs usually of the same or closely related breeds to identify, for example, single-nucleotide polymorphisms.[41] Given the particularity of Iberian breeds of pig and their local economic importance, it was not surprising that researchers in Spain would find it worth studying the genetics of these breeds in detail. When the work on distinct, regionally specific, economically important populations was breeding-oriented, it patterned a largely intra-national and often intra-institutional mode of collaboration, explaining UAB's lesser connectedness and centrality in this period—many publications lacked authors based in other countries or institutions. The intra-national co-authorships resulted in peripheral network positions, while the intra-institutional publishing left no co-authorship ties to be displayed with other nodes.

Once one gets into the mid- to late 2000s, however, the massive quantities of data deriving from the SGSP and subsequent full sequencing of specific pig breeds made criteria for inclusion in *Brief Notes* stricter, in line with earlier editorial shifts in other life science journals.[42] The papers reporting agriculturally relevant variants consequently tapered off and the publications associated with UAB started to focus more on genetic diversity, patterns of domestication, evolution, and phylogeny, expanding their range of target journals. These publications were characterized by multinational co-authorships, including with more marginal countries in the global south. Importantly, this shift to more extensive sequencing was conducted with some continuity of personnel from the preceding breeding-oriented work (as was also the case for

41. Sometimes this sequencing work was supplemented with association analyses to examine whether there was a correlation between the detected polymorphisms and variation in a given phenotypic trait or traits (e.g., in our dataset: PMIDs 15705744 and 16734680).

42. "Instructions to authors," *Animal Genetics* 39 (2008): 93–96. On how the mere description of sequence data became an insufficient milestone to be published in peer-reviewed journals and how this provides our corpus of publications with an added historiographical value, see Leng et al., "The Sequences" (n.10).

Wageningen).[43] UAB's latter work required, and used, the existence of reference sequence data (either in draft or more complete form), and therefore provides another indication of how the advent of a reference genome affected the kinds of research embodied in publications that are present in our dataset.

The distinctiveness of Iberian pigs is, in part, due to the relative lack of genetic inflow from pig breeds elsewhere, for example from Asia. Other European breeds show a considerable genetic flow from Asian populations, in some cases dating back to imports around 200 years ago.[44] The genetic idiosyncrasy of the animals the UAB researchers were working with conditioned the relatively inward-looking pattern of collaboration when breeding concerns were guiding intensive sequencing. They would have struggled to find partners beyond Spain who would want to work with the same Iberian breeds. The pattern became more outward-looking and international when extensive, systematic sequencing became paramount. For these studies, access to a wide array of expertise and geographically specific animals was required, entailing co-authorship with local institutions capable of obtaining and processing the DNA of these animals. This shows how the shift of emphasis from intensive to extensive sequencing, as well as the availability of inbred Iberian specimens and later the circulation of DNA samples from different pig populations for comparative purposes, underlies and shapes the co-authorship relationships in our network.

The latter co-authorship pattern of UAB exemplifies the network characteristics of our last institutional type and mode of research: institutions conducting systematic investigations in the extensive sequencing phase. UAB represents a shift from an intensive to an extensive form of sequencing within one institution, conditioned by the existence of the reference genome, the transformation of the research environment of genomics that it allowed, and the availability of a highly specific and idiosyncratic breed—the Iberian pig—that was used in distinct ways in intensive sequencing and comparative, extensive studies. The nodes corresponding to systematic institutions are often connected to subnetworks beyond the agricultural and hybrid institutions. Systematic institutions also exhibit more prevalent intercontinental co-authorships, especially those present in the betweenness centrality top twenty but not in the closeness centrality top twenty. An institution can occupy a place like this in the network due

43. Interview with Miguel Pérez Enciso, conducted by James Lowe over Skype, 28 September 2018.

44. Bin Yang, Leilei Cui, Miguel Perez-Enciso, Aleksei Traspov, Richard P. M. A. Crooijmans, Natalia Zinovieva, Lawrence B. Schook, et al., "Genome-wide SNP Data Unveils the Globalization of Domesticated Pigs," *Genetics Selection Evolution* 49 (2017): 71.

to the possession of particular kinds of pigs or certain capabilities such as expertise in the processing and interpretation of genomic data; these act as pull factors for collaboration. Many of them indeed suffered precarious financial support that, as mentioned above, led them to seek other funding opportunities and collaborate beyond their national borders.

UAB's transformation from predominantly intensive to extensive research and increasing pursuit of systematic inquiries over the period we examined shows the importance of the availability of the reference genome, as well as the plasticity and temporal contingency of sequence producers and users that the concept of bricolage foregrounds. Researchers at this institution—and oftentimes the same research groups—shifted over time from producing and reporting sequences concerning specific genes and polymorphisms to using the reference sequence as the basis for generating another type of data concerning genomic variation across breeds. In providing researchers with a thoroughgoing map of the sequence and genomic features, a species' reference genome is a resource that all new sequences can relate to and align against to inform the assembly and annotation of new sequences, and also to provide a platform for comparison and relation—or, in our analytic terms, for the extensification of sequencing practices. This comparability and capacity to act as an almost universal reference led to the growth of the systematic mode of research in our network from 2009 onward when the reference sequence became available.

Scholars have noted the importance of systematic practices in the collection and comparison of sequence information, genome mapping, whole-genome sequencing projects, and the exploitation of datasets generated by these projects.[45] Historian of science Jon Agar has wondered whether the advent of

45. On protein and DNA sequencing, see Bruno Strasser, *Collecting Experiments: Making Big Data Biology* (Chicago: University of Chicago Press, 2019); on taxonomy, see Dirk Stemerding and Stephen Hilgartner, "Means of Coordination in Making Biological Science: On the Mapping of Plants, Animals and Genes," in *Getting New Technologies Together: Studies in Making Sociotechnical Order*, eds. Cornelis Disco and Barend van der Meulen (New York: de Gruyter, 1998), 39–69; on genome mapping: National Academy of Sciences, "Mapping and Sequencing the Human Genome" (Washington, DC: The National Academies Press, 1988), 33; Stephen Hilgartner, *Reordering Life: Knowledge and Control in the Genomics Revolution* (Cambridge, MA: MIT Press, 2017), 31, 38; on constructing phylogenies based on comparable DNA sequence data across species: Edna Suárez-Díaz and Victor H. Anaya-Muñoz, "History, Objectivity, and the Construction of Molecular Phylogenies," *Studies in History and Philosophy of Biological and Biomedical Sciences* 39 (2008): 451–68. For an examination of the practices required to enable databases to perform these functions, see Sabina Leonelli, *Data-Centric Biology: A Philosophical Study* (Chicago: University of Chicago Press, 2016).

genomics heralds a new "golden age" of natural history.[46] In some cases, such as the Human Genome Project, systematic use of the data was not a key part of the original published analysis of the reference genome. It was, though, in the SGSP, which incorporated many pig geneticists who had maintained an interest in this angle dating back to the 1990s. By examining the production and use of the pig reference genome, we have shown how less well-known genomic endeavors reflected a much more specifically articulated deployment of systematics research, with application to questions concerning pig genetic diversity, patterns of domestication, and evolution borne in mind by many of the researchers involved in whole-genome sequencing.[47] They knew what they could draw from the reference genome, as well as what additional genomic resources they needed to produce in order to deepen their systematic understanding. This, in turn, relied on the reference genome as a fundamental scaffold. Bricolage thus reveals a cycle of entangled production and use of sequence data and associated tools and resources, which can be properly apprehended only by following how the processes unfold diachronically. Key to the bricolage of pig genomics and our ability as historians to uncover and conceptualize it is that the production and repurposing of tools and resources were often conducted by the same people—or at the very least the same communities. The cycles and practices represented by this bricolage were not merely created by a community, however; they were also the dynamic processes that helped to constitute and maintain the community itself.

The continuity of the research teams involved across the intensive and extensive phases at UAB enables us to comprehend how bricolage developed over time in our network. Systematic research was therefore another manifestation of the bricolage of pig genomics, one that received considerable impetus and resources from the determination of the reference sequence. Systematic

46. Jon Agar, *Science in the Twentieth Century and Beyond* (Cambridge, UK: Polity Press, 2012), 464–65.

47. In this paper, we use Human Genome Project to refer to both the initiative sponsored by the National Institutes of Health and US Department of Energy (1990 onward) and its subsequent internationalization up to the publication of the reference human genome sequence. In the special issue paper dealing with human genomics, earlier and later human genome projects are distinguished with the acronyms US-HGP and HGP: García-Sancho et al., "The Human Genome Project" (n.13). Human genomics and metagenomics initiatives conducted after the determination of the human reference sequence have produced data with systematic applications in mind, analogous to the model of pig genomics, see www.sanger.ac.uk/collaboration/25-genomes-for-25-years; Diana Marco, ed., *Metagenomics: Theory, Methods and Applications* (Norfolk, UK: Caister Academic Press, 2010).

institutions repurposed the reference genome and, rather than using it for agricultural and breeding goals, addressed other issues such as evolution, phylogeny, and genomic diversity across breeds and populations of *Sus scrofa*. This systematic research, in turn, fed back into the reference genome by providing key information on interbreed sequence variability.

The circularity in production and use that our bricolage framework reveals points to the central importance of the reference sequence, but not as the end product or culmination of a research endeavor. Approaching genomics as a continuous bricolage promotes an alternative view, one that considers the creation of reference sequences as open-ended; they constitute staging posts and inflection points in the history of genomic science rather than end points.

7. CONCLUSIONS

In this paper, we have portrayed pig genomics as a bricolage or repurposing of tools and materials, *bricolage* meaning both the process by which the tools and materials are repurposed and the outcomes of the repurposing process. Our bricolage framework, in line with the other papers of this special issue, has enabled us to further problematize the depiction of large-scale sequencing centers as the emblematic organizational model of genomics research. The bricoleurs that characterized pig genomics worked in different institutional forms and configurations, and their practices of bricolage differed from those that oriented around the comprehensive sequencing of whole genomes. Highlighting bricoleurs and bricolage in less prominent and more modestly funded initiatives than the Human Genome Project allows us to discern the open-endedness of sequencing data and the associated creation and deployment of tools and materials that circulate across communities.[48] This counters teleological interpretations of the reference sequence as an end point that marks the frontier between genomics and post-genomics research. Although not unique to genomics, the concept of bricolage—especially in its diachronic dimension—thickens the historiographical boundaries of genomics. In line with the other analytical tools presented in this special issue, bricolage enables us to portray the history of genomics as co-constitutive of other fields—such as

48. Bricolage and bricoleurs also existed in human genomics but are more difficult to detect due to the profound discontinuity between the communities conducting the whole-genome sequencing project and those using the resulting data for medical genetics research: García-Sancho et al., "The Human Genome Project" (n.13).

animal breeding and immunology, in the case of pig genomics—rather than being narrowly focused on the production of a reference sequence.[49]

During what we called the intensive phase of pig genomics, agriculturally inclined institutions contributed to both the production of the reference sequence and the characterization of genes and other genomic variants associated with livestock production traits. In pig genomics, in contrast to human genomics, the production of a reference sequence was not conceptualized as a separate endeavor to the aims of the agriculturally oriented research, or to involve an entirely different set of actors. It was, rather, a collaborative effort in which the sequencing undertaken by the Sanger Institute made use of genomic resources and tools (such as DNA libraries and radiation hybrid panels) developed by the agriculturally inclined institutions that coordinated the project and constituted some of the primary data users. Some of these agriculturally oriented institutions also contributed to the assembly and annotation of the resulting sequence. This paper has therefore expanded the conceptual reach of our earlier analysis of the SGSP.[50] We have shown that paying attention to the bricolage conducted by the institutions involved may not only widen our conception of sequencing practices but also thicken the historiography of genomics into the extensive sequencing phase and systematic research institutions.

The temporal entanglements between the production of a reference sequence and its uses are further shown during the extensive phase of pig genomics. The fact that some agriculturally inclined institutions that earlier conducted intensive sequencing—such as Wageningen University, the Roslin Institute, and UAB—were also engaged in the cross-breed comparisons characteristic of the systematic phase further documents the plasticity of the reference sequence and its capacity to stimulate different research agendas through evolving practices of bricolage. These practices unfolded before, during, and after the determination of the reference genome. Situating the reference genome in the context of the wider history of pig genomics without first reifying its centrality has thus enabled us to assess its consequences more richly. For example, the new kinds of extensive, systematic research facilitated by the

49. Returning to the wider theme of thriftiness and the circulation of tools and materials cited in note 16 above, we observe that the overt thriftiness of pig genomics may reflect the lower status of this field compared to other areas where thriftiness is more tacit, such as in human and yeast genomics; Mat Paskins, "Thrifty Science: Making the Most of Materials in the History of Experiment," *Ambix* 67, no. 2 (2020): 203–5.

50. Lowe, "Sequencing" (n.5).

advent of reference sequence data gave rise to more collaborative and international patterns of publication for institutions such as UAB. This institution's focus on a genetically distinct and geographically endemic Iberian breed also shows how access to material resources and specific expertise derived from working with them shaped practices of bricolage through the intensive and into the extensive phase.

We close with a reflection on the temporal dimensions concerning the production and use of sequence data. A reference sequence, once released, enables an extensification of sequence information. This engenders a perpetual cycle of augmenting the genomic resources and tools available, in which the distinction between producer and user is enacted only if we artificially freeze time. Rather than freezing time, in encompassing the totality of co-authored publications in our dataset, our network captured tokens of different aspects of bricolage, and thus enabled a temporal dimension to be apprehended. In attempting to understand the network patterns in conjunction with qualitative research, we had to reconstruct the historical background to the publications that underpinned those patterns, as well as their role in wider programs of research. Doing this helped us to appreciate the dynamic interrelation of sequence production and use, in which use requires production, and production requires use: in individual projects and programs, as well as in particular lineages of the production and repurposing of genomic tools and resources that we found.

Acknowledgments

See a full list of people and institutions whose support has been essential at the end of the introductory article of this special issue "The Sequences and the Sequencers: What Can a Mixed-Methods Approach Reveal about the History of Genomics?" The research and writing of this paper were sponsored by the "TRANSGENE: Medical Translation in the History of Modern Genomics" Starting Grant, funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant agreement No. 678757. For more details on the project, see https://transgene.sps.ed.ac.uk/.