

# DNA-Analyse und Vereinsgenealogie

## Erkenntnisse aus einer schwierigen Kooperation

Dass DNA-Genealogie ein wichtiges Thema für die bürgerwissenschaftliche Vereinsgenealogie ist, wird auf den ersten Blick niemanden überraschen. Etwa seit dem Jahr 2000 werden Analysen der DNA für die Ermittlung von Blutsverwandtschaftsbeziehungen angeboten, und es hat sich international eine lebhafte Szene von Aktiven entwickelt, die entsprechende Tests nutzen und auswerten. Was sich dort entwickelt, ist allerdings eine ganz andere Art von Genealogie als die archivgestützte: technischer, kommerzieller, stärker auf das 20. Jahrhundert fokussiert, auch: ausschließlich auf die biologischen, nicht auf die sozialen Verwandtschaftsbeziehungen bezogen. Hier zu berichten ist über das Experiment einer Kooperation zwischen Vereinsgenealogie und kommerzieller DNA-Genealogie, das der *Verein für Computergenealogie e. V.* (CompGen) in den vergangenen Jahren unternommen hat. In diesem Projekt sollte eine möglichst fein aufgegliederte regionale Kartierung genetischer Herkunft innerhalb Deutschlands erarbeitet werden. Die fachlichen Ergebnisse – so viel sei vorweggenommen – waren begrenzt; mehr als eine grobe Aufteilung des Untersuchungsgebiets in drei Unterregionen wurde nicht erreicht. Aufschlussreich für ein Verständnis dessen, was Genealogie in unserer Gesellschaft ist und sein kann, war das Projekt aber in hohem Maße.

CompGen hat ein großes Tätigkeitsfeld, für das der Vereinsname inzwischen eigentlich nur ein recht eng gewähltes Stichwort bietet: Nicht nur die Genealogie im Sinne von Ahnenforschung interessiert die Mitglieder und Aktiven, sondern auch das, was inzwischen als Familiengeschichtsforschung bezeichnet wird. Gemeint ist damit die Erkundung sozialer Beziehungen und Lebensumstände in der Geschichte, mit Blick auf personenbezogene Informationen nicht nur zur Abstammung, sondern auch zum Besitz von Häusern oder Autos, zu Kriegserfahrungen oder Wohnverhältnissen, zu den Gesichtern der Menschen früherer Generationen auf Familienfotos oder auch zu mündlich oder schriftlich überlieferten Erinnerungen. Zudem geht es dem Verein nicht allein um niederschwellige Hinweise zum fachlich sinnvollen Benutzen des Computers, sondern auch um den Einsatz von aktuellen Informatikmethoden in der bürgerwissenschaftlichen Geschichtsforschung, um Digitalisierung von Quellen, Texterkennung, *crowdsourcing* (also das gemeinsame Erfassen von Quellen im Netz) und um Datenbankmodelle, die anschlussfähig zu modernen offen zugänglichen Wissensspeichern sind. Es ist also nicht überraschend, dass ein Verein, der auf dem Feld zwischen dem Massen-Hobby Ahnenforschung und modernen wissenschaftlichen Methoden tätig ist, auch offen ist für die Herausforderung, die für die Genealogie durch das Aufkommen der



kommerziellen DNA-Genealogie in den beiden Varianten der *genetic genealogy* (Genetische Genealogie) und des *genetic ancestry testing* (Herkunftsanalyse) entstanden ist.

## Das Projekt „Eine Familie – die Deutschen“

Im Jahr 2017 sind CompGen und die britische Firma LivingDNA eine Zusammenarbeit eingegangen, um eine verbesserte auf DNA-Analysen beruhende Herkunftsschätzung für den deutschsprachigen Raum zu entwickeln. Das Projekt „Eine Familie – die Deutschen“<sup>1</sup> war im Vergleich zu anderen CompGen-Projekten immer untypisch, weil es nicht ums Mitmachen ging wie bei den auf dem Dateneintragssystem<sup>2</sup> beruhenden *crowdsourcing*-Projekten, dem Sammeln von Grabstein-Fotos<sup>3</sup> und vielen anderen. Stattdessen – für uns als nichtkommerziellen Verein sehr ungewöhnlich – ging es um den Erwerb von vergünstigten Test-Kits, die dann von LivingDNA ausgewertet wurden. Teilnehmen konnten alle Personen, deren Großeltern nicht weiter als 80 km auseinander geboren waren, weil man versuchte, eine für die genetische Zusammensetzung der Bevölkerung vor den Migrationsprozessen des 20. Jahrhunderts aussagekräftige Stichprobe zu generieren. Mittlerweile ist die Analyse der Test-Kits abgeschlossen, sodass – hier aus der subjektiven Sicht des zuständigen Vorstands-Resorts Citizen Science – die Erfahrungen mit dem Projekt abschließend eingeordnet werden können.<sup>4</sup>

Über die Jahre wurden vermutlich etwa 200 Test-Kits bei dieser Aktion verkauft, etwa 110 (sowie – genau weiß man es nicht – wohl ein paar Dutzend aus anderen Datenquellen) wurden letztlich bei der Analyse berücksichtigt. Aufseiten von CompGen überließ der Projektinitiator die Projektleitung bald einem engagierten Mitglied, das allerdings seinerseits nach einiger Zeit ausschied, sodass das Projekt in letzter Zeit verwaist war. Mehr als Kommunikationsarbeit, teils erklärend, teils im Sinne einer Verteidigung der DNA-Genealogie-Firmen gegen sehr unterschiedlich motivierte Kritik, sah die Projektleitung in diesem ungewöhnlichen Projekt nicht als ihre Aufgabe; die DNA-Genealogie wurde in Mailinglisten, in einem eigenen Themenportal im genealogischen Wiki<sup>5</sup> sowie außerhalb des Verantwortungsbereichs von CompGen in einer Facebook-Gruppe thematisiert. Eine Kommunikation über Analysestrategien und inhaltliche Interpretationen fand nicht statt; hier bestand eher die Erwartung, dass die Firma diese zu gegebener Zeit liefern werde.

---

1 Timo Kracke, DNA Projekt: Eine Familie – Die Deutschen, <https://www.compgen.de/2017/04/dna-projekt-eine-familie-die-deutschen/> (9.10.2021).

2 <https://des.genealogy.net/> (9.10.2021).

3 <https://grabsteine.genealogy.net/> (9.10.2021).

4 Der Verfasser war seit 2018 als Beisitzer für Citizen Science und Beziehungen zur akademischen Forschung im Vorstand des *Vereins für Computergenealogie*, seit 2019 als zweiter Vorsitzender. Er ist seit November 2021 erster Vorsitzender. Er war nicht in die Entwicklung oder Leitung des Projekts involviert, sondern hatte auf Vorstandsebene eine Zuständigkeit für wissenschaftliche Kooperationsvorhaben, zu der auch die Evaluation von Ergebnissen gehört.

5 <http://genwiki.genealogy.net/Portal:DNA-Genealogie> (9.10.2021).

## DNA-Genealogie: Praktiken und Grenzen

Das Anliegen des Projekts, die Herkunftsanalyse, macht einen praktisch für die Genealogie relativ unwichtigen, aber in der Öffentlichkeits- und Werbewirksamkeit eben doch sehr sichtbaren Teil der DNA-Genealogie aus. Grob gesagt: DNA-Genealogie besteht aus Herkunftsanalyse und DNA-*matching* – und letzteres ist der wichtigere Teil. Dabei nutzt man DNA wie ein Stückchen ‚Quellentext im Körper‘, also prinzipiell nicht anders als andere Quellen: Man möchte konkrete Verbindungen finden und überprüfen. Es gibt Situationen, wo es zu dieser Methode überhaupt keine Alternative gibt. Andererseits kann ein DNA-*matching* keine mit der klassischen Kirchenbuch-Genealogie vergleichbare zeitliche Reichweite erreichen, schon weil die meisten DNA-Spuren nach ein paar Generationen verschwunden sind.

Als Statistiker gesprochen: Ein Maß für den Anteil gemeinsamen Genoms zwischen Verwandten ist die *coancestry*, deren Erwartungswert z.B. bei Geschwistern bei  $(\frac{1}{2})^2$ , also einem Viertel liegt. Sie ist nicht fest determiniert, sondern folgt einer Zufallsverteilung, bei Geschwistern mit einem Konfidenzintervall zwischen etwa 20 und 30 Prozent. Daraus folgt, dass es aufgrund des gegebenen Streubereichs noch zu praktisch 100 Prozent sicher ist, bei Geschwistern solche Genomabschnitte vorzufinden, die auf die gemeinsamen Eltern und nicht auf die sicherlich auch vorhandenen gemeinsamen Vorfahren vor vielen Generationen zurückgehen. Anders gesagt: Wenn zwei Vollgeschwister genealogische DNA-Tests machen, werden diese sicher auf eine enge biologische Verwandtschaft hindeuten. Geht man aber 150 Jahre zurück, also die etwa fünf Generationen bis zur Gründung des Kaiserreichs, und wieder vor in die Gegenwart zu den Cousins oder Cousinen vierten Grades, dann fällt der Erwartungswert für die *coancestry* auf  $(\frac{1}{2})^{10}$  – das ist dann nur ein Promille. Mit einer schon recht hohen Wahrscheinlichkeit von über 20 Prozent besteht mit der Cousine vierten Grades also nur eine homöopathisch verdünnte Blutsverwandtschaft, aber keine auf die gemeinsamen Urururgroßeltern zurückzuführende genetische Verwandtschaft. Geht man noch ein oder zwei Generationen zurück, dann wird die (in diesem Sinne verstandene) genetische Verwandtschaft, bezogen auf ihren Anteil an den vorhandenen Blutsverwandten, immer mehr zur *relativen* Ausnahme, auch wenn die *absoluten* Trefferzahlen noch in beeindruckendem Maße steigen.<sup>6</sup> DNA-Genealog\*innen versuchen nun oft, gegen diesen statistischen Effekt einer Verdünnung des Wissens anzuarbeiten und die Chance auf Treffer wieder etwas zu erhöhen, indem sie möglichst viele potentielle oder tatsächliche Verwandte dazu bewegen, ebenfalls Tests zu machen. Auch hieraus erklärt sich der in der Szene verbreitete Eifer, das eigene Verfahren zu propagieren und gegen Kritik zu verteidigen.

Gründung des Kaiserreichs – dieser Zeitpunkt trennt für die Genealogie in Deutschland zwei Zeiträume: Für die Zeit vor 1874 (Einführung der Standesämter) benutzt die Familienforschung in Deutschland in der Regel Kirchenbücher. Für die Zeit danach ist die Familienüberlieferung wichtiger (weil Kirchenbücher lange Zeit nur bis zum Jahr 1874 zugänglich waren). Bildlich gesprochen: DNA als ‚Quellentext‘ enthält für die Zeit von etwa 1870 bis in die Gegenwart eine Art tendenzielle Komplettdokumentation der Abstammungsverhältnisse. Für die Jahrhunderte davor steht diese tendenzielle Komplettdokumentation (von zufälli-

---

6 Siehe hierzu Doug Speed/David J. Balding, Relatedness in the Post-Genomic Era: Is it Still Useful?, in: Nature Review Genetics 16 (2015), 33–44, DOI: 10.1038/nrg3821, bes. 35, Tabelle 1. Ich danke Mark G. Thomas (University College London) für Hinweise.

gen Verlusten und Migrationseffekten abgesehen) dagegen in den Kirchenbüchern, während im ‚Quellentext‘ DNA sehr systematisch und gründlich der Radiergummi sein Unwesen treibt. Diejenigen Verwandtschaftsbeziehungen, für deren Nachweis ich Kirchenbücher brauche, könnte ich mit DNA-*matching* allein nur gelegentlich feststellen. Als Hobby wird DNA-Genealogie daher meist in Form eines Nachdoppelns der ‚Papierforschung‘ betrieben. DNA-*matching* kann aber durchaus auch Lücken im Wissen über die Familiengeschichte seit der Kaiserzeit füllen – ein Wissen, das in vielen Familien mündlich oder auch durch die Ahnenforschung früherer Generationen gut überliefert ist, in anderen Familien aber vergessen wurde.

## Herkunftsanalyse und die Frage „Wer bin ich?“

Herkunftsanalyse ist – im Gegensatz zum *matching* – technisch sehr wenig stabil und zugleich missverständlich. Die verbreiteten, von Anbieter zu Anbieter deutlich schwankenden Prozentangaben suggerieren, dass die eigenen Vorfahren zu einem bestimmten Anteil aus bestimmten Regionen wie *Northeast Europe* oder *Baltics* stammen, als ob man eine komplette Ahnentafel rückwärts bis zu einem gedachten Stichjahr statistisch nach Regionen ausgewertet hätte. Tatsächlich bietet Herkunftsanalyse nur auf einer Skala von 0 bis 100 umgerechnete und daher in Prozentwerten darstellbare Ähnlichkeiten der eigenen DNA zu jener der heutigen Bevölkerung dieser Regionen (ermittelt mithilfe derjenigen Proben, die sich in den konkreten Datenbanken der Firmen finden). Regionale Verteilung in der Ahnentafel und Ähnlichkeit zu heutigen regionalen DNA-Mustern sind mitnichten dasselbe: Nicht nur seit der Industriellen Revolution, deren Anfänge immerhin auch schon eine Handvoll Generationen zurückreichen, sondern auch in der vorindustriellen Gesellschaft war Migration über kurze, mittlere oder auch längere Distanzen weit verbreitet. Es ist also nicht überraschend, wenn eine Person – sagen wir – Menschen aus dem heutigen Litauen genetisch mit ein paar „Prozentpunkten“ ähnelt, ohne auch nur einen einzigen litauischen Vorfahren zu haben.

Für diejenigen, die sich – wie es in der genealogischen Forschung typisch ist – für die Familiengeschichte Einzelner interessieren, ist Herkunftsanalyse also nur als eine Art Wegweiser sinnvoll, und auch nur dann, wenn die angezeigten Anteile einerseits erheblich sind und andererseits auch auf Regionen deuten, in denen man plausiblerweise mit Archivquellen weiterforschen kann. Eine solche Wegweiserfunktion haben in ähnlicher Weise – nur stärker differenziert – auch Namensverbreitungskarten, die daher auch in der genealogischen Praxis eine erhebliche Rolle spielen.<sup>7</sup> Nur wenn man die Perspektive vom einzelnen Datenpunkt hin zur Population wechselt, wird es – grob gesagt – möglich, aus der stärkeren oder schwächeren genetischen Ähnlichkeit Rückschlüsse auf räumliche Tendenzen und die ungefähre zeitliche Erstreckung von Ereignissen (etwa Wanderungen) zu ziehen.<sup>8</sup>

Während sich diese Rückschlüsse auf Strukturen von Populationen beziehen, ist Herkunftsanalyse andererseits aber auch sehr stark mit Bedeutung für das Individuum aufgeladen: Leicht glauben Einzelne, auf diesem Wege etwas über die persönliche ‚Ethnizität‘ zu

7 CompGen bietet ein entsprechendes Tool unter <https://nvk.genealogy.net/map> (9.10.2021) an.

8 Das Verfahren wurde 2014 entwickelt von: Garrett Hellenthal u.a., A Genetic Atlas of Human Admixture History, in: Science 343 (2014), 747–751, DOI: 10.1126/science.1243518.

erfahren. Während die Pfade zu den Blutsverwandten mithilfe von atDNA nur über kurze historische Zeiträume hinweg einigermaßen vollständig rekonstruierbar sind, bieten YDNA und mtDNA auch sehr langfristige Verbindungen zu „Adam“ und „Eva“ genannten steinzeitlichen Urvätern und Urmüttern – allerdings nur in der als Haplogruppe bezeichneten rein männlichen oder rein weiblichen Linie.<sup>9</sup> Mit Herkunfts- wie Haplogruppenanalyse eröffnen sich Gedankenspiele für Antworten auf die Frage „Wer bin ich?“ mit hoher zeitlicher Reichweite. Die Versuchung ist nämlich groß, DNA nicht nur als ‚Quellentext im Körper‘ zu verstehen, sondern als Essenz dessen, wie – um hier Formulierungen aus der DNA-Genealogie aufzugreifen – wir „Menschen aus verschiedenen Regionen der Welt“ uns selbst „mehr oder weniger stark voneinander unterscheiden“.<sup>10</sup> Wer sind die Gruppierungen, über die man hier sagt, dass sie sich „unterscheiden“?

In der amerikanischen Kultur, aus der die kommerzielle DNA-Genealogie überwiegend stammt, spielt ‚Ethnizität‘ eine große Rolle, also ob man schwarz, weiß, *Hispanic* oder *Native American* ist. Man trägt diese Information in amtliche Formulare ein, man hat einen bevorzugten Zugang zu bestimmten Ressourcen aufgrund der ‚Ethnizität‘, man streitet sich über Fehlzuordnungen, man wird aber auch entsprechend beäugt und eingeordnet. *Ethnicity* und *race* werden weitgehend als Synonyme verstanden, und Vorstellungen, die an der biologischen Abstammung und am körperlichen Erscheinungsbild anknüpfen, spielen hierfür eine zentrale Rolle.<sup>11</sup>

In Deutschland ist ‚Ethnizität‘ heute als politischer Begriff in diesem Sinne von etwas geringerer Bedeutung. Gewiss werden Deutsche mit nicht ‚biodeutsch‘ wirkenden Gesichtszügen auch bei uns noch oft anders angeschaut. Aber viele persönliche Eigen- und Erbschaften – Familiengeschichten –, die uns im Leben umtreiben, sind mit einem Blick auf den Körper ganz und gar nicht zu erkennen. Bin ich ein Nichtjude oder ein Jude, eine Ost- oder eine Westdeutsche? Erkenne ich in den Geschichten, die mir mündlich überliefert wurden, eine ‚kulturprotestantische‘ oder eine katholisch-,provinzielle‘ Familienkultur? Stamme ich aus Russland, aus dem Emsland, aus Polen, Bayern oder Siebenbürgen – und welche ‚Etiketten‘ verbindet meine Umwelt damit, dass meine Familie von hier oder dort kommt?

---

9 G. David Poznik u.a., Sequencing Y Chromosomes Resolves Discrepancy in Time to Common Ancestor of Males Versus Females, in: Science 341 (2013), 562–565, DOI: 10.1126/science.1237619.

10 Zitat aus: <http://genwiki.genealogy.net/DNA-Herkunftsanalyse> (9.10.2021), verfasst von einem Aktiven der Hobby-DNA-Genealogie. Interessant ist hier der Sprachgebrauch, der sich zur Erklärung der Funktionsweise von DNA-Herkunftsanalyse auf Eigenschaften „der Menschen“ und nicht auf solche der von ihnen über die Generationen weitergegebenen körpereigenen Stoffe bezieht.

11 Die amerikanische Begrifflichkeit (zu der es eine unermessliche Forschungsliteratur gibt) lässt sich an den Fragen nach *race* und *ethnicity* im seit 1790 alle zehn Jahre erhobenen Zensus gut nachvollziehen. Begriffe wie „Hautfarbe“ (*color*) und „Herkunft“ (*ancestry*) spielen oder spielten dort wechselnde Rollen. Aktuell wird *race* vom Bureau of the Census als Kategorie sozialer Selbstidentifikation definiert, während die American Anthropological Association darauf verweist, dass der Rassenbegriff als biologischer Begriff verstanden wird und Vorstellungen objektiver Unterschiede zwischen Menschen transportiert, die sachlich unzutreffend sind: American Anthropological Association, Statement on Race, May 1998, <https://www.americananthro.org/ConnectWithAAA/Content.aspx?ItemNumber=2583> (9.10.2021).

## DNA-Genealogie und Wissenschaft

Mit diesen Fragen nach Kultur und Identität im Hintergrund eröffnet sich durch die DNA-Analyse potentiell doch etwas mehr, als dass sie bloß eine Ergänzung der personenbezogenen Forschung um ein paar nur hochtechnisch auslesbare ‚Quellentexte im Körper‘ bietet und damit als ein zusätzlicher Wegweiser für die ‚Papierforschung‘ dient. Fragt man nach der Stellung der DNA-Genealogie im Kontext des Wissenschaftssystems, fallen zwei interessante Praktiken auf, die beide miteinander verbunden sind.

Die eine besteht darin, dass sich hier – wie schon während der bürgerlichen Neuerfindung der Genealogie um 1900 – die Gelegenheit zu bieten scheint, Wissen über relevante Eigenschaften von Menschen zu gewinnen und im Kontext der Herkunft zu erklären. In den entsprechenden Medien tummeln sich rechte Freigeister, die sich für die biologischen, körperlichen Aspekte der Herkunft vor allem im Bemühen darum interessieren, positive Identitäten zu schaffen – etwa wenn sie bestimmte Charaktereigenschaften (z.B. Intelligenz oder die Neigung zu ADHS) auf bestimmte, ins Indogermanische zurückreichende Haplogruppen zurückführen.<sup>12</sup> Das geschieht auf durchaus wissenschaftsförmige Art – womit das zweite Problem angesprochen ist. In der DNA-Genealogie versucht man – nicht anders als in der ‚Papier‘-Genealogie – nachvollziehbar zu arbeiten; und das Streben nach Nachvollziehbarkeit macht ja einen Kern von Wissenschaftlichkeit aus. Kann Wissenschaft dort neu erfunden werden, wo sich eine Gruppe von Kund\*innen einzelner Firmen untereinander austauscht und die wichtigsten Informationsquellen aus der Firmenkommunikation stammen? Gibt es tragfähige kommunikative Brücken zwischen wissenschaftlicher (Archäo-)Humangenetik und Hobbyforschung? Kann Überprüfbarkeit ohne ihre Institutionalisierung funktionieren?

DNA-Forschung beruht auf hochtechnischer Naturwissenschaft. Wenn Fachfremde wissenschaftliche Forschungen näher verfolgen, besonders im Grenzbereich zwischen Natur- und historischen Sozialwissenschaften, sammeln sie rasch erhebliche Methodenkenntnisse an. Das ist im Bereich der Bürgerwissenschaften, auch der Genealogie, ganz allgemein der Fall. Als der Verfasser dieses Beitrags vor etwa 35 Jahren das erste Mal im Archiv mit vor Anstrengung rotem Kopf versuchte, eine Quelle für die geplante Dissertation zu entziffern, bat er einen weißköpfigen Genealogen um Lesehilfe und bekam – über die Hilfe hinaus – zur Antwort: „Das ist doch eigentlich eine sehr klare Handschrift...“. Wer sich die Wissenschaft zum Hobby macht, kann also oft mehr, viel mehr, als ein ‚Studierter‘ des jeweiligen Faches.

Die schwierige Frage ist die, ob das, was man als Vielwisseur oder Vielkönnlerin in der Hobby-Forschung oder auch in der Diskussion über die Forschung tut, dadurch selbst zur Wissenschaft wird – nicht nur als wissenschaftlich gemeint, sondern als vom Wissenschaftssystem akzeptiertes, für ‚gut‘ befundenes Forschungsergebnis. Der Test hierfür, also für die heikle Frage, ob das jetzt nun ob seiner Schwierigkeit tatsächlich auch (gute) Wissenschaft ist, liegt in formalen Verfahren der Qualitätskontrolle. Anders gesagt: Wissenschaftlichkeit erkenne ich nicht daran, dass ein Text besonders schwierig zu verstehen ist und dadurch Respekt einflößt, sondern dass es Wege gibt, ihn ganz respektlos auseinanderzunehmen und

---

12 So meint der Wissenschaftsblogger Ingo Bading, aus der Erforschung der *ancient DNA* Erkenntnisse über die Herkunft der „großen, blonden, blauäugigen Rohmilchverdauer mit hoher angeborener Intelligenz, mit ADHS und großen thymotischen Energien“ zu gewinnen: Ingo Bading, Was macht uns Europäer genetisch so einzigartig?, <http://studgendettsch.blogspot.com/2017/07/die-trichterbecher-leute-standen.html> (9.10.2021).

zu entschlüsseln, und dass andere Personen – Gutachter\*innen, Herausgeber\*innen, Prüfungsgremien usw. – ihn auch bereits entschlüsseln und verstehen konnten.

Bekanntlich können solche Verfahren auch fehlerhafte Ergebnisse erbringen – absoluten Schutz vor Irrwegen bieten sie nicht. Letztlich sorgen aber nur sie dafür, dass die Ergebnisse eine Anschlussverwendung in professionellen und akademisch organisierten Bereichen finden. Bei Forschungsdaten ist das ganz offensichtlich: Datensammlungen brauchen Verfahren, die ihre Nutzbarkeit nach den FAIR-Regeln (*findable, accessible, interoperable, re-usable*) garantieren. Dass wir in *crowdsourcing*-Projekten Quellen erfassen, nützt wenig, wenn die Ergebnisse nicht als *Linked Open Data* im Netz abfragbar sind. Das in der Genealogie-Szene entstandene Geschichtliche Ortsverzeichnis, das historische Siedlungs- und Verwaltungsstrukturen besser modelliert als jedes andere Geodatenystem, wird erst dann zum wissenschaftlichen Standard, wenn die in der Wissenschaft zuständigen Organisationen das auch bemerken und bestätigen.

Ähnliches gilt für Forschungstexte: Wer wissen will, wie seriös die Arbeiten anderer sind, schaut auf die Publikationsorte. YouTube-Kanäle, Facebook-Auftritte, selbstgestaltete Webseiten oder auch die Werbematerialien von Firmen sind nicht zitierfähig. Wissenschaft funktioniert nicht ohne institutionalisierte Qualitätskontrolle – und zwar besonders dann, wenn sie anspruchsvoll und technisch nicht gerade niederschwellig ist. Beispiele des Scheiterns können im harmloseren Fall darin bestehen, dass jemand ein sich wissenschaftlich gerierendes Zeitschriftenprojekt in Form einer Website mit PDFs – ohne Verankerung an akademischen Institutionen – aufbaut, das dann auch bald wieder eingestellt wird.<sup>13</sup> Im weniger harmlosen, an die oben skizzierte Forschung nach menschlichen Eigenschaften und Unterschieden anschließenden Fall werden Texte geschrieben, die durchaus kenntnisreich und detailliert anmuten, deren Erscheinungsort in der vom Verfassungsschutz beobachteten Zeitschrift *Reconquista*<sup>14</sup> jedoch darauf hindeutet, dass ein Inselepublikum von bereits einschlägig Überzeugten angesprochen werden soll.

Positivbeispiele, die punktuell auch in der institutionalisierten Wissenschaft wahrgenommen werden, entstehen, wenn Nachvollziehbarkeit durch Aufbau von Datenbeständen systematisch gesichert und unterstützt wird. Solche Datenbestände bestehen etwa in Sammlungen von *matching*-Ergebnissen, die den (labortechnisch in *centiMorgan* gemessenen) Abstand zwischen den Genomen zweier Personen in Beziehung setzen zum auf dem Wege der klassischen Genealogie ermittelten Verwandtschaftsgrad.<sup>15</sup> Dadurch werden nicht nur die theoretisch oder aufgrund von Simulationsrechnungen erwartbaren Übereinstimmungen, sondern vor allem deren empirisch durch die Kombination von Zufallsprozessen und Messfehlern entstehenden Streuungen erkennbar.

---

13 Journal of Genetic Genealogy, letzte Ausgabe Herbst 2016: <https://jogg.info/> (9.10.2021).

14 Hier finden sich mehrere Beiträge des DNA-Genealogen „Max Kaltmeister“, u.a. im Blog „Archäogenetik & Frühgeschichte“: <http://www.reconquista.jetzt/index.php?id=852> (9.10.2021).

15 Blaine T. Bettinger, The Shared cM Project: A Demonstration of the Power of Citizen Science, in: Journal of Genetic Genealogy 8/1 (2016), 38–42; neuere zusammenfassende Statistiken (jedoch keine Forschungsdaten) hierzu auch unter <https://thegeneticgenealogist.com/2020/03/27/version-4-0-march-2020-update-to-the-shared-cm-project/> (26.11.2021). Einschränkend muss allerdings bemerkt werden, dass die ausgewerteten Daten, also die selbstberichteten DNA-Testergebnisse, eine starke Verzerrung aufweisen, da Null-Ergebnisse von den Freiwilligen nicht berichtet wurden und in der statistischen Analyse dieses Problem nicht angemessen adressiert wurde.

Ein weiterer Problembereich besteht in der Vereinbarkeit von Bürgerwissenschaft, Open Science und laborgestützter Forschung. Die klassische Kirchenbuch-Genealogie spielte sich in einem Dreieck ab: Hobby – Pfarrarchiv – Wissenschaft. An die Stelle des Pfarrarchivs stellen sich in letzter Zeit mehr und mehr kommerzielle Dienste, die den Zugang zu den Quellen erleichtern. Im Fall der DNA-Genealogie ist die Rolle der kommerziellen Dienste dabei weit mehr als die von Vermittlern. Sie haben ein Monopol auf Techniken und Methoden, die sonst nur in wissenschaftlichen Labors genutzt werden. Die Frage, ob diese Firmen seriös sind, ist jeweils offen. Im wissenschaftlichen Sinne trifft das dann zu, wenn sie ihre Verfahren und die gesammelten Daten offenlegen. Dass Firmen unseriös agieren, ist gerade im Bereich der DNA-Genealogie durchaus eine reale Möglichkeit; berüchtigt war etwa die Firma BritainsDNA, die auf Kritik erst mit großer Aufgeregtheit, dann mit juristischen Mitteln reagierte und schließlich aufgelöst wurde. Die fließende Grenze zwischen seriöser genetischer Herkunftsanalyse (in der Begrifflichkeit des Teams um Mark Thomas: *genetic ancestry*) und unseriöser Spekulation (*genetic astrology*) eröffnet einen großen, auch kommerziellen Spielraum.<sup>16</sup>

## Partner und Erträge

All das könnte einen gemeinnützigen Verein, der an wissenschaftlicher Reputation orientiert und ohnehin an der kommerziellen Genealogie nicht aktiv beteiligt ist, dazu bewegen, von der ganzen Sache die Finger zu lassen. Der (damalige) CompGen-Vorstand ist einen anderen Weg gegangen, nämlich den, mit einem möglichst wissenschaftsnahen Anbieter zusammenzuarbeiten. LivingDNA erschien damals vor allem wegen ihres Anspruchs, die Methoden des auf über 6.200 Fällen beruhenden genetischen Projekts *People of the British Isles* (PoBI) zu replizieren, als (relativ) wissenschaftlich.<sup>17</sup> In der Kooperationsvereinbarung wurde angestrebt – und fairerweise muss man sagen: „angestrebt“ heißt bei wissenschaftlichen Projekten nicht „versprochen“ –, publizierbare Ergebnisse zu erarbeiten, die grundsätzlich denen des PoBI-Projektes vergleichbar sein sollten. Diese Ergebnisse sollten – so LivingDNA in der Kooperationsvereinbarung – von „unseren Wissenschaftlern“ stammen. Gemeint war die Arbeitsgruppe GenSci, zu der auch drei am PoBI-Projekt beteiligte Wissenschaftler gehör(t)en (Garrett Hellenthal, Daniel Falush und Simon Myers) und die mit LivingDNA kooperiert(e). Die Verbindung zu diesen „unseren“ Wissenschaftlern scheint eher lose zu sein, es sind soweit erkennbar keine Beschäftigten von LivingDNA; Hellenthal firmiert etwa als Berater des Unternehmens. LivingDNA hat seine wissenschaftlichen Kontakte nach außen teilweise recht euphorisch kommuniziert.<sup>18</sup> Aber sie existieren.

---

16 Debbie A. Kennett u.a., The Rise and Fall of BritainsDNA: A Tale of Misleading Claims, Media Manipulation and Threats to Academic Freedom, in: *Genealogy* 2/4, Artikelnummer 47, 1–26, DOI: 10.3390/genealogy2040047.

17 Stephen Leslie u.a., The Fine-Scale Genetic Structure of the British Population, in: *Nature* 519 (2015), 309–314, DOI: 10.1038/nature14230.

18 Selbstbeschreibung nach <https://cruwys.blogspot.com/2016/09/living-dna-new-genetic-ancestry-test.html> (9.10.2021): „Living DNA is a collaboration of over 100 world-leading scientists, academic researchers and genetic experts“.

Man sollte sich LivingDNA freilich nicht als die kommerzielle Nebenaktivität eines wissenschaftlichen Projektteams vorstellen, es handelt sich eher um ein paar Menschen, die verschiedene Inputs auf kreative Art und Weise miteinander verbinden und damit (legitimerweise) auch Geld verdienen – ein klassisches Beispiel dafür, wie Unternehmertum funktioniert. Einer dieser Inputs besteht in wissenschaftlichem Know-how, den im PoBI-Umfeld entwickelten Methoden. Ein anderer ist die konkrete Labortechnik, die vom Luxemburger Labormulti Eurofins stammt, welcher auch die Kooperation mit CompGen initiiert hat. Ein dritter besteht in den Kundenbeziehungen. Diese sollten durch die Kooperation mit CompGen verbessert werden – allerdings ist es LivingDNA nie gelungen, sich durch kontinuierliche und transparente Kundenkommunikation eine positive Reputation zu erarbeiten, wahrscheinlich aus Gründen der fehlenden Versiertheit in kontinentaleuropäischen Sprachen. Der vierte Input sind dann die von der genealogischen Community eingesammelten Test-Kits mit Proben und die daraus gewonnenen Daten. Grundsätzlich können – so die Nutzungsbedingungen – die Daten auch für Forschungszwecke genutzt werden, die sich im Lauf der Zeit ändern. Sie bilden also ein auch für nichtgenealogische Forschungsinteressen möglicherweise hochinteressantes Datenkorpus.

Aus den Inputs – darunter eben den Datensammlungen – stellt das Unternehmen Outputs, Produkte, her und vermarktet diese. LivingDNA ist in zwei Bereichen tätig: *Ancestry Kits* (also Herkunftsanalysen und *matching*) und *Wellbeing Kits* (also DNA-Tests, aus denen Empfehlungen zu Ernährung, Nahrungsergänzung und Sport abgeleitet werden). Die Muttergesellschaft DNA-Worldwide bietet darüber hinaus auch Einwanderungs- und Drogentests für Behörden an. Interessant ist das Anwendungsfeld *Wellbeing*, das mit den genealogischen Analysen zusammen vermarktet wird. Hier verfügt LivingDNA über vertiefte Kontakte durch eine persönliche Verbindung<sup>19</sup> der Firmengründer\*innen Nicholson und Morden zum ebenfalls in Frome (Somerset) ansässigen, von Serge Benhayon angeführten esoterischen Gesundheits- und Ernährungskult *Universal Medicine*, auf dessen Doktrin *Way of the Livingness* schon der Firmenname verweist.<sup>20</sup>

Von den Outputs eines Unternehmens zu unterscheiden sind die wissenschaftlichen Publikationen, die auf den auch vom Unternehmen verwendeten Techniken beruhen. Aus dem Umfeld von PoBI und GenSci bzw. mit entsprechenden Methoden (etwa den Verfahren *ChromoPainter* und *FineSTRUCTURE*) sind in letzter Zeit z.B. Untersuchungen zur genetischen Herkunftsstruktur von Finnland und den Niederlanden erschienen, mit der grundsätzlichen Fragestellung, ob aus der DNA Erkenntnisse über Migrationsmuster der Vergangenheit gewonnen werden können. Die Studie zu Finnland stellt z.B. aufgrund von fast 2.400 DNA-Proben einen langfristigen Effekt des Vertrags von Nöteborg (1323) fest, der einen (ungefähren) Grenzverlauf zwischen den Herrschaftsrechten Schwedens und der Republik Nowgorod festschrieb und damit einen Einfluss auf das Siedlungsverhalten späterer Epochen nahm.<sup>21</sup> Die Studie zu den Niederlanden zeigt anhand von Proben von etwa 1.400 Personen,

---

19 Margaret O'Brien, Who Owns LivingDNA (Explained), <https://www.dataminingdna.com/who-owns-living-dna/> (9.10.2021); David Nicholson, Universal Medicine – The Facts of My Experience, <https://web.archive.org/web/20200228135845/https://truthaboutuniversalmedicine.com/2016/02/22/universal-medicine-the-facts-of-my-experience/> (9.10.2021).

20 [https://en.wikipedia.org/wiki/Living\\_DNA](https://en.wikipedia.org/wiki/Living_DNA) (9.10.2021).

21 Sini Kerminen u.a., Fine-Scale Genetic Structure in Finland, in: *G3 Genes|Genomes|Genetics* 7 (2017), 3459–3468, DOI: 10.1534/g3.117.300217.

dass Migration langfristig überwiegend Flussläufen in Ost-West-Richtung folgte.<sup>22</sup> Beide Studien sind nicht so sehr als inhaltlicher Beitrag zu aktuell diskutierten Forschungsfragen bedeutsam als in methodischer Hinsicht.

Die relativ wenigen aus der Kooperation in Deutschland gewonnenen Daten bieten keine Basis für Analysen wie in Großbritannien, den Niederlanden oder Finnland. Als Wissenschaftler, der sich für Migrationsgeschichte interessiert, kann ich das bedauern. Aber für die Forschungen der meisten Vereinsmitglieder stehen nicht Fragen wie jene im Vordergrund, wie wichtig Flussläufe oder feudale Herrschaftsverträge für Umzüge, Hofgründungen und Heiraten im Hoch- und Spätmittelalter oder in der Frühneuzeit waren. Eher geht es im Hobby um die Frage, wohin mich meine persönliche Forschung führt, wenn ich die Menschen anschau, die mich geprägt haben, und die, die diese Menschen geprägt haben, und die davor und die, die auch mit diesen Menschen in Zusammenleben und Begegnungen in Kontakt waren – also um ein potentiell unendlich weites Netz aus persönlichen Verbindungen.<sup>23</sup> Hierfür brauche ich Wegweiser, wenn die kirchenbuch- oder urkundenbasierte Forschung nicht weiterführt. Diese sollen Räume und Zeiten benennen, in denen die Suche vielleicht lohnt.

Worin also besteht das Ergebnis des Kooperationsprojekts? Im Frühjahr 2020 präsentierte LivingDNA eine Karte, nach der Deutschland und einige seiner Nachbarländer in drei Gebiete zu unterteilen sind: „nordwestgermanisch“ (Teile Dänemarks, Niederlande, westelbisches Nord- und Mitteldeutschland), „nordostgermanisch“ (zwischen Elbe und Oder/Neiße; anders als mit CompGen vereinbart hat LivingDNA die bis Mitte des 20. Jahrhunderts zu Deutschland gehörigen Gebiete weggelassen), „südgermanisch“ (ein bunter Mix vom frankophonen Unterwallis bis nach Österreich, wenn auch ohne Niederösterreich, Wien und Burgenland, wieder einschließlich Westtschechiens, dann einer ungefähr geraden, aber freihändig gezeichneten Grenze bis zum Rhein folgend, dann ein Stück den Rhein entlang, das Großherzogtum Luxemburg einschließend, die belgische Provinz Luxemburg aber abschließend bis einschließlich Flandern).<sup>24</sup>

Angesichts der wenigen zugrundeliegenden Datenpunkte sind die von Hand gezogenen Grenzen höchstens als hypothetisch zu bezeichnen, von der merkwürdigen Benennung der Gebiete als „germanisch“ ganz abgesehen. Ein Beitrag zur wissenschaftlichen Forschung ist dieses Ergebnis also offensichtlich nicht. Zu lesen ist es bestenfalls so: Wenn ich bei LivingDNA einen Test mache und mir z.B. ein nennenswerter Anteil „nordostgermanisch“ angezeigt wird, dann lohnt sich möglicherweise die weitere Recherche in deutschen Archiven östlich der Elbe. Die praktische Relevanz dieser Herkunftsanalyse ist nicht groß.

Aus der Sicht des Vorstandsressorts Citizen Science kann festgehalten werden: Erstens, wissenschaftliche Projekte sind Versuche, und nicht jeder Versuch gelingt. Zweitens, auf Forschungsk Kooperationen mit Unternehmen sollte sich ein dem Open-Data-Grundsatz verpflichteter bürgerwissenschaftlicher Verein nur dann einlassen, wenn diese hinsichtlich der Offenheit von Forschungsdaten und Forschungsverfahren eine hohe Transparenz zeigen,

---

22 Ross P. Byrne u.a., Dutch Population Structure across Space, Time and GWAS Design, in: Nature Communications 11 (2020), Artikelnummer 4556, 1–11, DOI: 10.1038/s41467-020-18418-4.

23 Elisabeth Timm, Reverenz und Referenz. Zwei Weisen der populären Genealogie seit dem 19. Jahrhundert und ein neuer genealogischer Universalismus?, in: Christine Fertig/Margareth Lanzinger (Hg.), Beziehungen – Vernetzungen – Konflikte. Perspektiven Historischer Verwandtschaftsforschung, Köln 2016, 209–231.

24 Die Karte wird verkleinert auf <https://livingdna.com/eu/european-ancestry> (9.10.2021) gezeigt, detaillierter u.a. auf <https://european-origins.com/2020/04/23/living-dna-ancestry-test-review/> (9.10.2021).

auch wenn – oder vielleicht: besonders wenn – sich das Unternehmen im Allgemeinen als ganz besonders wissenschaftlich präsentiert. Denn: Drittens, nicht immer ist die Grenze zwischen Wissenschaft und grobem Unsinn von vornherein klar zu erkennen. Nicht für Wissenschaftler\*innen, und auch nicht für Nicht-Wissenschaftler\*innen. Viertens, DNA-Genealogie kann missverstanden werden als Beitrag zur Frage, wer ‚wir‘ als ganze Population ‚eigentlich‘ sind, und Antworten aus tiefster, lange vor der Kirchenbuchführung liegender Vergangenheit suggerieren. Das hat mit Genealogie nichts zu tun (mit Geschichtswissenschaft auch nicht). Fünftens, DNA-Tests machen Quellen im Körper lesbar. Es mag Einwände gegen diese Tests geben, die ich hier nicht angesprochen habe (etwa datenschutzrechtliche, politische, ethische) – aber als Historikerinnen und Historiker neigen wir tendenziell immer dazu, Quellen, die lesbar sind, auch zu lesen. Sechstens, Forschung kann von Wegweisern profitieren. CompGen sollte sich daher mit dem Thema Namensforschung weiter beschäftigen.

Das Projekt ist – nicht mit dem erhofften Erfolg, aber doch auch nicht ganz ohne Ertrag – beendet.