

The Changing Text of Acts: A Phylogenetic Approach^{*}

Pasi Hyytiäinen, University of Helsinki

Abstract: New Testament textual critics have long maintained that the earliest textual tradition of the Acts of the Apostles is bipolar, transmitted in two early textual forms. This conviction is now being challenged, with recent studies suggesting that the tradition is more complex than the two-text concept had proposed. How should we then approach Acts? This article approaches Acts from an evolutionary point of view, applying phylogenetic methods to its manuscripts. Scholars have been using computer-assisted phylogenetic methods for years to produce trees and networks that describe the relationships among manuscripts within a textual tradition. These methods were originally developed for evolutionary biology, but studies have shown that they can also be applied to manuscript traditions. Here, these methods are applied to selected manuscripts to test their applicability, since Acts has never been subjected to such a study. Chapter 5 of Acts is used as a test case to demonstrate how phylogenetic analysis can be conducted. The preliminary results point to a complex set of relationships among manuscripts, concurring with recent hypotheses about the complexity of the tradition. At the same time, however, these methods do recognize the two early textual groups of Acts. While it seems that the tradition in Acts 5 is too complex to be fitted into a single tree, a network is capable of depicting the complexity of the tradition.

Keywords: textual criticism, text-type theory, computer-assisted stemmatology, Acts 5, genealogical methods, phylogenetics, phylomemetics, networks

By the early eighteenth century, New Testament critics and natural scientists, facing the same problem of how to describe the relationships within their growing databases, produced the same solution: hierarchical classification.¹ Johann Albrecht Bengel (1687–1752) envisioned a grand project of classifying the New Testament manuscripts into nations, tribes, families, and companies.² This forward-looking vision was not fully carried out by Bengel or by any other subsequent New Testament textual critic,³ but it did lead to a threefold scheme for describing the manuscript families—two ancient textual groups (Alexandrian and Western) and a later

^{*} I wish to express my deep gratitude to Professor Chris Howe for all his advice and the anonymous reviewer whose comments helped to improve and clarify the text.

¹ Yii-Jan Lin, *The Erotic Life of Manuscripts: New Testament Textual Criticism and the Biological Sciences* (Oxford: Oxford University Press, 2016), 27–31.

² Carl Linnaeus also used these so-called nested hierarchies in his classification of the living world into kingdoms, classes, orders, genera and species. See Johann Bengel, *Apparatus criticus ad Novum Testamentum, Briseos sacrae compendium, limam supplementum acfructum exhibens* (Tubingae, Sumtibus Io. Georgii Cottae, 1734), 387; Carl Linnaeus, *Systema naturae in quo naturae regna tria, secundum classes, ordines, genera, species, systematice proponuntur* (Stockholmiae apud Gottfr. Kiesewetter, 1735).

³ Bengel distinguished two great groups or nations of manuscripts, the Asiatic and the African, the latter being, it seems, subdivided into two tribes. See Bengel, *Apparatus criticus*, 429–30; *Defensio Novi Testamenti Graeci Tubingae anno 1734 editi* (Lugd. Bat. Wishoff, 1737), 47.

one (Byzantine)—forming the basis for the development of the text-type theory that scholars still use today.⁴ The threefold scheme has, at first glance, taken a somewhat different form when applied to the text of Acts of the Apostles, commonly referred to as the two-text theory of Acts. This theory maintains that there are two early textual forms or streams in Acts, that is, Alexandrian and Western, though at the same time it is acknowledged that the emerging Byzantine text played a crucial role in the later development of the text. Hence, the theory of two texts is, in fact, the text-type concept in Acts as far as the earliest phases of the textual transmission are concerned.⁵

At the core of the two-text concept of Acts lie two prominent codices, the fourth-century Codex Vaticanus (henceforth B03) and Codex Bezae (henceforth D05), a Greco-Latin codex copied in the early fifth century CE. Their widely differing texts have been the main source of the two-text concept, with all new manuscripts—or at least the older ones—of Acts having been classified with B03 as Alexandrian or with D05 as Western.⁶ In recent years, Eldon Jay Epp has proposed two clarifications for the text-type theory. First, one should abstain from using the term *text-type* because it is “too rigid, too tightly circumscribed and definitive”; one should prefer the term *textual cluster* instead. Second, the traditional geographical designations of these groups (Alexandrian, Western, Byzantine) should be changed to more neutral terms: A-Textual Cluster (Byzantine), B-Textual Cluster (Alexandrian), and D-Textual Cluster (Western).⁷

Despite these developments, the two-text concept of Acts is now being challenged as a whole, with recent studies suggesting that the complexity of the tradition in Acts exceeds the early two-text hypothesis. When the fifth-century papyrus manuscript P127 was first published in 2009, containing fragments of Acts 10–12 and 15–17, the editors David Parker and Stuart Pickering noticed that it is impossible to assign P127 to one or the other of the two early textual forms of Acts, raising questions about the validity of the two-text concept.⁸ Most recently, Georg Gäbel and Klaus Wachtel have argued that there is no evidence indicating the existence of a single early Western text (or D-text) of Acts; thus, the notion of such a text should be aban-

⁴ See the influential formulation of the text-type theory in Brooke Foss Westcott and Fenton John Anthony Hort, *The New Testament in the Original Greek: Introduction and Appendix to the Text Revised* (New York: Harper, 1882), 119–35; Eldon Jay Epp, “Textual Clusters: Their Past and Future in New Testament Textual Criticism,” in *The Text of the New Testament in Contemporary Research: Essays on the Status Questionis*, ed. Bart D. Ehrman and Michael W. Holmes, 2nd ed., NTTSD 42 (Leiden: Brill 2013), 519–77.

⁵ That is, one finds two early textual forms and one later also in Acts. Epp argues on behalf of two opposing texts in Acts, but at the same time for the influence of the emerging Byzantine text on the Western textual group. David Parker has also referred to two competing forms of text, while speaking of Byzantine text as representing the later text of Acts. See David C. Parker, *Introduction to the New Testament Manuscripts and Their Texts* (Cambridge: Cambridge University Press, 2008), 173–4; David C. Parker and Stuart R. Pickering, eds., “4968. Acta Apostolorum 10–12. 15–17,” in *Graeco-Roman Memoirs* 95 (London: Egypt Exploration Society; British Academy 2009), 1–45, esp. 8; Epp, “Textual Clusters,” 560; “Text-Critical Witnesses and Methodology for Isolating a Distinctive D-Text in Acts,” *NovT* 59 (2017): 225–96, esp. 230–32.

⁶ Parker, *Introduction*, 293–301; Parker and Pickering, “4968. Acta Apostolorum,” 3–6.

⁷ There is also a Caesarean textual cluster (C-Cluster), but its existence has been questioned from many quarters. See Larry Hurtado, *Text-Critical Methodology and the Pre-Caesarean Text: Codex W in the Gospel of Mark*, SD 43 (Grand Rapids, MI: Eerdmans, 1981), 88; Eldon Jay Epp, “The Significance of the Papyri for Determining the Nature of the New Testament Text in the Second Century: A Dynamic View of Textual Transmission,” in *Perspectives On New Testament Textual Criticism: Collected Essays 1962–2004* (Leiden: Brill 2005), 345–81, esp. 359–63; “Methodology for Isolating,” 226.

⁸ Parker and Pickering “4968. Acta Apostolorum,” 6–8.

done once and for all.⁹ On the other hand, the Alexandrian text rests on a handful of witnesses, which is, as Parker put it, “claiming a great deal from very little.”¹⁰ Thus, a question presents itself: if we are not dealing here with a two-text problem, how should we then approach Acts?

To tackle this question, we turn to the natural sciences, particularly to evolutionary biology. This article approaches Acts from an evolutionary point of view, applying computer-assisted phylogenetic methods originally designed for the analysis of biological evolution.¹¹ Previous studies have already demonstrated that such advanced programs can be successfully applied to textual data.¹² Indeed, the sorts of variations found in the texts of the New Testament (e.g., contamination, coincidental rise of variants)¹³ have close parallels to those described in the study of evolutionary biology (e.g., genetic recombination, convergent evolution, etc.).¹⁴ Scientists make use of different programs, such as PAUP*,¹⁵ to draw phylogenetic trees from DNA sequences—that is, taxonomical tree diagrams that describe evolutionary relationships among organisms. Texts can also be inputted into these programs as if DNA sequences, enabling interdisciplinary collaboration between the fields of evolutionary biology and textual criticism.¹⁶

It should be noted that an evolutionary approach to texts is not a novel standpoint in New Testament textual criticism. Already in 1924, Burnett Streeter portrayed the development of local texts in a Darwinian manner, suggesting that a text acquires unique features as it adapts, so to speak, to its local environment.¹⁷ Epp has similarly described the New Testament as a “living, breathing organism reflecting and reacting to its social and theological environment.”¹⁸ In

⁹ Georg Gäbel, “Western Text, D-Text Cluster, Bezan Trajectory, Or What Else?—A Preliminary Study,” in *Novum Testamentum Graecum: Editio Critica Maior; The Acts of the Apostles* [hereafter ECM], ed. Holger Strutwolf et al. (Stuttgart: Deutsche Bibelgesellschaft, 2017), 3/3: 83–136, esp. 133–34; Klaus Wachtel, “On the Relationship of the ‘Western Text’ and the Byzantine Tradition of Acts—A Plea Against the Text Type Concept,” in ECM 3/3: 137–48, esp. 147.

¹⁰ Parker, *Introduction*, 171.

¹¹ This approach is sometimes called *phylomemetics*, but here the terms evolutionary, phylogenetic and phylomemetic analysis are used interchangeably. See Christopher J. Howe and Heather F. Windram, “Phylomemetics—Evolutionary Analysis Beyond the Gene,” *PLoS Biology* 9 (2011): 1–5.

¹² Christopher J. Howe et al., “Parallels between Stemmatology and Phylogenetics,” in *Studies in Stemmatology II*, ed. Pieter van Reenen, August den Hollander and Margot van Mulken, vol. 2 (Amsterdam: Benjamins 2004), 3–11; Caroline Macé and Philippe V. Baret, “Why Phylogenetic Methods Work: The Theory of Evolution and Textual Criticism,” in *The Evolution of Texts: Confronting Stemmatological and Genetical Methods*, ed. Caroline Macé et al. (Pisa, Rome: Istituti Editoriali e Poligrafici Internazionali 2006), 89–108.

¹³ Contamination occurs when scribes use more than one exemplar when copying a text. A coincidental rise of variants when the same reading occurs in distantly related texts.

¹⁴ Genetic recombination is the process whereby two different copies of a gene come together to produce a hybrid, called a recombinant. Convergent evolution occurs when the same change arises independently in different lineages. See Howe et al., “Parallels,” 7–9.

¹⁵ David L. Swofford, “PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4,” Sinauer Associates, Sunderland, Massachusetts (2003), doi: 10.1111/j.0014-3820.2002.tb00191.x.

¹⁶ The text is encoded into a program by using different letters or numbers to indicate different variations. Alternatively, binary sequences—that is, a combination of 0s and 1s—can be used to encode variations. See Christopher J. Howe et al., “Manuscript Evolution,” *Trends in Genetics* 17 (2001): 147–52.

¹⁷ Burnett Hillman Streeter, *The Four Gospels: A Study of Origins* (London: Macmillan, 1924), 35–9.

¹⁸ Eldon Jay Epp, “Textual Criticism in the Exegesis of the New Testament,” in *Perspectives on New Testament Textual Criticism: Collected Essays 1962–2004* (Leiden: Brill 2005), 461–95, esp. 478.

fact, there seems to be increasing interest in the evolutionary approach, with the phylogenetic methods having been applied to different New Testament texts in the past few years.¹⁹

However, Acts has never been subjected to evolutionary analysis.²⁰ We thus apply this approach to selected manuscripts to test the phylogenetic methods and evaluate the results against the recent hypothesis about the complexity of the tradition. Chapter 5 of Acts is used here as a test case to demonstrate how phylogenetic analysis can be conducted. The cluster terminology developed by Epp is adopted in this survey for a couple of reasons. First, we agree with Epp that *text-type* is an overly rigid term to describe the textual groups, at least in Acts. *Clusters*, on the other hand, need not necessarily bear strong and clearly marked similarity throughout the manuscripts included or exhibit a large close-knit group of manuscripts at the centers, as Epp has put it.²¹ Second, the cluster terminology is also used in phylogenetics to signify a group of taxa that share the same branch; that is, they are related to one another and have a common ancestor.²² This is not far from textual clusters, since textual agreement between members of a given cluster is often explained as their having some common textual basis; that is, they have common descent. In other words, the cluster terminology can be seen as another link between phylogenetics and textual criticism, and hence the traditional textual groups are referred to here as A-Cluster (Byzantine), B-Cluster (Alexandrian), and D-Cluster (Western).

Preliminary Remarks: Phylogenetics and Stemmatology

Phylogenetic methods often use trees as models to depict the history of the entities studied,²³ placing the approach in proximity with stemmatology, which goes back to the pioneering work of Karl Lachmann (1793–1851).²⁴ Lachmann's method involves constructing family trees or stemmata to describe the relationships between the surviving manuscripts of a given text. These stemmata are drawn using common (or indicative) errors—that is, if manuscripts contain the same errors and such errors do not have a high probability of occurring independent-

¹⁹ Matthew Spencer, Klaus Wachtel and Christopher J. Howe, "The Greek Vorlage of the Syra Harclensis: A Comparative Study on Method in Exploring Textual Genealogy," *TC* 7 (2002): <http://rosetta.reltech.org/TC/v07/SWH2002>; Stephen C. Carlson, *The Text of Galatians and Its History* (Tübingen: Mohr Siebeck, 2015); Jac D. Perrin, *Family 13 in St. John's Gospel: A Computer Assisted Phylogenetic Analysis* (Leiden: Brill, 2018); Andrew C. Edmondson, "An Analysis of the Coherence-Based Genealogical Method Using Phylogenetics" (PhD diss., University of Birmingham, 2018).

²⁰ However, techniques used in phylogenetics, such as hierarchical clustering, have been applied to classify manuscripts of Acts. See Gerald J. Donker, *Text of the Apostolos in Athanasius of Alexandria* (Atlanta: Society of Biblical Literature, 2011), 216–26, 296–304.

²¹ Note that Epp himself has argued that the two textual streams in Acts are the text-type concept in Acts. Hence, the relation between "cluster" and "text-type" remains somewhat vague. "Textual Clusters," 553, 560.

²² Anne-Mieke Vandamme, "Basic Concepts of Molecular Evolution," in *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, ed. Philippe Lemey, Marco Salemi, and Anne-Mieke Vandamme, 2nd ed. (Cambridge, UK: Cambridge University Press, 2009), 3–29, esp. 20.

²³ This is not always the case, for some phylogenetic methods (such as NeighborNet) produce networks rather than branching trees. Howe and Windram, "Phylomemetics," 3.

²⁴ Lachmann founded the principles of the method, which were worked out by several successive critics. See, e.g., Giorgio Pasquali, *Storia della tradizione e critica del testo*, 2nd ed. (Firenze, 1952); Paul Maas, *Textual Criticism* (London: Oxford University Press, 1958).

ly, the manuscripts are taken to be of common descent.²⁵ This logic interestingly aligns with the most parsimonious principle in phylogenetics, according to which a common character is inherited directly from a common ancestor.²⁶

There are, however, some fundamental differences between these approaches. The phylogenetic analysis of manuscripts is not restricted to a few indicative errors that have been considered to be of stemmatic importance, but it also utilizes variants that may have arisen by coincidence. Although Lachmannians see this decision as a mistake, adding noise to the data, studies have shown that categorizing variants according to significance has little effect on the structure of the tree as long as the data are generally consistent.²⁷ On the other hand, if the variants support many contradictory trees, then it is difficult to handle the data by either manual or computer-assisted methods.²⁸

While the stemma or the family tree as used in the textual scholarship is always oriented—that is, the tree graph has a root or an archetype, which is the most recent common ancestor of all the manuscripts in the stemma²⁹—phylogenetic methods are not designed to indicate the position of the root unless further information is given. Initially, an unrooted tree is produced, displaying the relatedness of the manuscripts without making assumptions about ancestry. However, a critic can root the tree using whatever information is at his or her disposal and then use this rooted tree as a stemma to identify ancestral readings.³⁰

Some critics question the phylogenetic approach since these methods are designed to produce bifurcating trees; that is, every internal node or split in the tree has a maximum of two descendants. In the manuscript traditions, however, it was common for several copies to be made from a single exemplar.³¹ Many phylogenetic methods are indeed constructed to produce bifurcating trees, but the resulting tree can also contain multifurcations. This may occur when the statistical robustness of a tree is tested by using bootstrapping, for instance.³² If there is little statistical support for a particular bifurcation, the tree will be shown as a multifurcation.³³ Also, the result of the phylogenetic analysis need not be a tree at all. Phylogenetic methods can also produce networks.

²⁵ Maas, *Textual Criticism*, 42–49; Ben Salemans, “Building Stemmas with the Computer in a Cladistic, Neo-Lachmannian Way, The Case of Fourteen Text Versions of Lanseloet van Denemerken” (PhD diss., Katholieke Universiteit, Nijmegen, 2000), 64–71.

²⁶ Vandamme, “Basic Concepts,” 27.

²⁷ See, e.g., Matthew Spencer et al., “The Effects of Weighting Kinds of Variants,” in *Studies in Stemmatology II*, 227–40.

²⁸ Christopher J. Howe, Ruth Connolly, and Heather F. Windram, “Responding to Criticisms of Phylogenetic Methods in Stemmatology,” *Studies in English Literature 1500–1900*, 52 (2012): 51–67, esp. 60.

²⁹ Martin L. West, *Textual Criticism and Editorial Technique Applicable to Greek and Latin Texts* (Stuttgart: Teubner, 1973), 32.

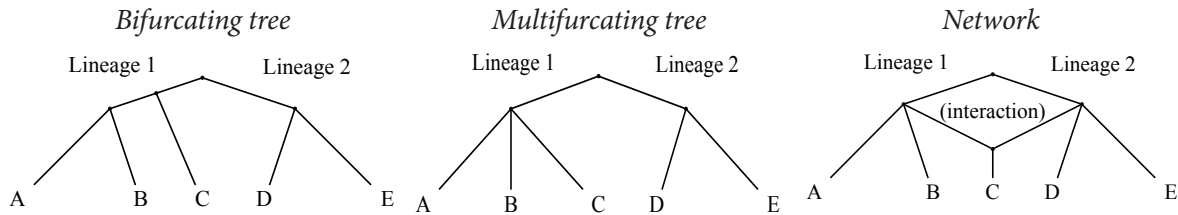
³⁰ Robert J. O’Hara and Peter Robinson, “Computer-Assisted Methods of Stemmatic Analysis,” in *Canterbury Tales Project Occasional Papers*, ed. Norman F. Blake and Peter M. W. Robinson, vol. 1 (Oxford: The Office for Humanities Communication, 1993), 53–74; Howe, Connolly and Windram, “Responding to Criticisms,” 62.

³¹ Harold Love, “The Work in Transmission and Its Recovery,” *Shakespeare Studies* 32 (2004): 73–80, esp. 77.

³² See below, p. 17.

³³ Howe, Connolly and Windram, “Responding to Criticisms,” 63–64.

Figure 1. The difference between trees and a network



The difference between different trees and a network becomes clear in the simple example above, where each split represents the hypothetical ancestors (internal node) and letters written in capitals as extant manuscripts. In the bifurcating tree, all internal nodes split in two, whereas the multifurcating tree allows the manuscripts A, B, and C to descend from one ancestral node, called *polytomy* in phylogenetics. Bifurcating and multifurcating trees create separate lineages, but networks allow cycles to be formed.³⁴ Hence, a network is able to depict interactions between different lineages, which is a useful feature when analyzing contaminated traditions, such as Acts.

The phylogenetic approach also differs from the Coherence-Based Genealogical Method (CBGM), which has recently been adopted in New Testament textual criticism. This groundbreaking toolbox was developed at the Institut für Neutestamentliche Textforschung (INTF) by Gerd Mink.³⁵ CBGM creates trees or stemmata through computational analysis with the assistance of philological judgments by textual critics. The starting point of the method is to construct a local stemma (i.e., a hypothesis about the relationship of different readings at one place of variation) for every place of variation in a given text. Here, a critic evaluates which disagreeing variants are prior and which are posterior. These prior and posterior readings are calculated for every manuscript. The more prior readings a particular manuscript contains, the earlier the textual state it represents. This data is then combined with statistical information about the relatedness (based on agreements between witnesses) of any given pair of witnesses. CBGM can thus be seen as a hybrid of computer analysis and traditional textual critical methods.³⁶

In other words, critics using CBGM have to scrutinize every place of variation and differentiate prior readings (ancestral) from posterior ones. In the phylogenetic approach, on the other hand, every text is encoded with a string of characters, designating all places of variations and variants without any assumptions about their relationships. This leads to a computer-generated unrooted tree (or network), which the critic can first work with before determining the direction of descent and the root of the tree.³⁷ Furthermore, while CBGM is a closed system—that is, the underlying data cannot be changed without administrative rights—every step of the phylogenetic methods can be managed, evaluated, and reassessed. It should also be noted that there is not a single computer-assisted phylogenetic method but a wide variety of ways to infer trees, making it possible to compare the resulting trees and perhaps gain a more coherent picture of the manuscript tradition under investigation.

³⁴ Vincent Moulton and Katharina T. Huber, “Split Networks: A Tool for Exploring Complex Evolutionary Relationships in Molecular Data,” in Lemey, Salemi, and Vandamme, *Phylogenetic Handbook*, 631–53.

³⁵ See the general introduction to the method by Gerd Mink in “CBGM Presentation” (presented at the Münster Colloquium on the Textual History of the Greek New Testament, Münster, August 3–6, 2008).

³⁶ Lin, *Erotic Life*, 124–25.

³⁷ O’Hara and Robinson, “Computer-Assisted Methods,” 61–62.

Lachmannian stemmata and phylogenetic trees also differ from CBGM in their relation to hyparchetypes or subancestors (internal nodes). These stand for manuscripts that have been lost in time, since actual exemplars from which a given manuscript was copied rarely ever surface. CBGM, however, excludes these subancestors from its stemmata, since CBGM does not—as Lachmannian stemmata and phylogenetic trees do—try to indicate which manuscripts were copied from which. Instead, it tries to exhibit the relations among texts (texts are differentiated from their physical carriers—viz., manuscripts), depicting “a structure that integrates states of text and must not be confused with an outline of actual historical processes or a stemma of manuscripts.”³⁸

In summary, the phylogenetic approach does not use any one type of variant but rather an overview, which is in turn used to infer a tree (bifurcating or multifurcating) or a network. These methods do not indicate the position of the root, for the textual data inputted to the programs only contain information about the places of variation and variants without any prior assumptions about ancestry. The critic can first work with the unrooted tree or network before determining the root and the direction of ancestry based on all his or her knowledge of and expertise in the manuscript tradition in question. If the underlying data is consistent, that is, not supporting several contradictory tree topologies (tree structures), the result of the analysis can be a rooted tree, which can be treated as a stemma of manuscripts—that is, a hypothesis of the history of the manuscript tradition.

Selecting the Data

The Institut für Neutestamentliche Textforschung (INTF) currently has 528 witnesses of Acts catalogued, of which 274 contain Acts 5.³⁹ Due to the limited time and scope of this survey, we have chosen to include only a representative sample. The importance of the prominent majuscules, B03 and D05, with the witnesses agreeing with one or the other, cannot be underestimated here. Chapter 5 of Acts was chosen for this analysis, since it contains several shorter and longer places of variation where B03 and D05 disagree, and since these places of variation are supported by other witnesses, hence giving a good approximation of the textual challenges that Acts contains as a whole. This chapter also allows us to include some relevant witnesses in the phylogenetic analysis, which would be unavailable in some other chapters, such as papyrus P45 and majuscules C04 and 0189. The length of the chapter (forty-two verses) also enables a relevant comparison of the witnesses.

The B- and D-Clusters (Alexandrian and Western) are, however, very different in nature in Acts 5. The B-Cluster consists of a larger group of witnesses containing papyri P45 and P74, majuscules 801, A02, B03, and 0189 (showing a high degree of agreement with B03), and minuscule 1175. These are considered as the primary⁴⁰ witnesses to the B-Cluster in this survey,

³⁸ According to the logic of CBGM, texts can relate in ways manuscripts cannot—e.g., a young manuscript can contain a very old text and may then be ancestor to the text of a much older manuscript. Spencer, Wachtel and Howe, “The Greek Vorlage of the Syra Harclensis,” para. 14; Klaus Wachtel, “The Coherence Method and History,” *TC* 20 (2015): 1–6, esp. 1–2, <http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-history.pdf>; Tommy Wasserman and Peter J. Gurry, *A New Approach to Textual Criticism: An Introduction to the Coherence-Based Genealogical Method* (Atlanta: SBL Press, 2017), 24.

³⁹ See <http://ntvmr.uni-muenster.de/liste>.

⁴⁰ That is, they are manuscripts that represent the B-Cluster in the purest manner, even though it must be underlined that all manuscripts are mixtures of more than one textual cluster. This mix-

due to their high agreement rates in Acts 5 (above 90 percent, except for B03 and P45 with 85 percent agreement),⁴¹ even though A02 and 1175 are often seen as secondary witnesses to this textual cluster.⁴² In addition, the secondary level of this cluster includes majuscule C04 and, with reservation, minuscule 1409, both located in Category II by the Alands.⁴³

The D-Cluster, on the other hand, faces a very different situation. D05 is the only Greek manuscript in Acts 5 that can be labeled as belonging to the D-Cluster. Here, a cluster is identified using the well-known calculation of E. C. Colwell that a cluster is a group of manuscripts that agree more than 70 percent of the time and is separated by a gap of about 10 percent from its neighbors.⁴⁴ None of the other Greek manuscripts seem to meet these conditions. Other witnesses to the D-Cluster are Latin, Syriac, and Coptic manuscripts. To counterbalance this deficiency, four important Latin manuscripts were translated into Greek, two of which have been categorized by Epp as the primary witnesses to the D-Cluster, namely, the Latin side of D05 (d5) and Fleury Palimpsest (h55) and two secondary ones, Codex Gigas (gig51) and Codex Perpinianensis (p54).⁴⁵ Here, however, these five witnesses are treated as the core or the primary level of the D-Cluster due to their agreement rates, which exceed 75 up to 85 percent (86 between D05 and d5 and 85 percent between gig51 and p54).⁴⁶ The secondary witnesses to this cluster include E08 and 1884.⁴⁷

This is not an ideal situation, for translating Latin manuscripts into Greek entails several difficulties. Word order, for instance, cannot be translated directly from one language to the next. Also, Greek uses definite articles, whereas there are no articles in Latin; hence, it is occasionally difficult to identify which Greek reading the Latin text actually represents or what the Greek text is behind the Latin. For these reasons, we have relied on critical editions, particularly ECM, and other available literature to decide which readings to follow.⁴⁸ An example will clarify how the translation is conducted, which is, as will become clear, not so much a

ture increases when moving to the secondary witnesses to a given textual cluster. Concerning the textual mixture, see Kurt Aland and Barbara Aland, *The Text of the New Testament: An Introduction to the Critical Editions and to the Theory and Practice of Modern Textual Criticism*, 2nd rev. ed. (Grand Rapids, MI: Eerdmans 1995), 70, 335.

⁴¹ Minuscule 81 would belong to this group, but unfortunately it has lacunae in 4:8–7:17.

⁴² See, e.g., Bruce M. Metzger and Bart D. Ehrman, *The Text of the New Testament: Its Transmission, Corruption and Restoration*, 4th ed. (Oxford: Oxford University Press, 2005), 313.

⁴³ Minuscule 1409 contains a high proportion of A-Cluster readings in Acts 5; hence, it is not entirely sure whether it qualifies as a witness to the B-Cluster. See Aland and Aland, *The Text of the New Testament*, 109, 129, 134. Concerning minuscule 33, see below p. 11.

⁴⁴ This is the definition of text-types by Colwell, but the same definition was used by Epp when speaking of textual clusters. E. C. Colwell, "The Quantitative Relationships between MS. Text-types," in *Biblical and Patristic Studies in Memory of Robert Pierce Casey*, ed. J. N. Birdsall and R. W. Thomson (Freiburg: Herder, 1963), 29; Epp, "Significance of the Papyri," 372.

⁴⁵ Epp, "Methodology for Isolating," 235–39.

⁴⁶ These values are taken from a distance matrix generated by PAUP*. See pp. 21–22.

⁴⁷ This is despite the fact that E08 and 1884 are far apart from the primary witnesses, with their agreement values being below 70 percent.

⁴⁸ Sources include, *inter alia*, James Hardy Ropes, *The Acts of the Apostles: The Text of Acts*, vol. 3.1 of *The Beginnings of Christianity*, ed. F. J. Foakes Jackson and Kirsopp Lake (London: Macmillan, 1926), 44–57; Bruce M. Metzger, *A Textual Commentary on the Greek New Testament: A Companion Volume to the United Bible Societies' Greek New Testament*, 2nd ed. (Stuttgart: Deutsche Bibelgesellschaft, 1994); Jenny Read-Heimerdinger and Josep Rius-Camps, *Acts 1.1–5.42: Jerusalem*, vol. 1 of *The Message of Acts in Codex Bezae: A Comparison with the Alexandrian Tradition*, JSNTSup 257 (London: T&T Clark, 2004), 294–365; ECM 3/1.1: 131–72.

translation per se but a textual-critical analysis to determine which Latin reading corresponds with which particular Greek variant.

In the first verses of Acts 5, Peter learns that Ananias and Sapphira kept some of the money they got by selling their farm, and he accuses Ananias of lying to God. When Ananias hears this, he falls down on the ground:

Table 1. Translating Acts 5:5a

	A	B	C	D
B03	ἀκούων	δὲ ὁ Ἀνανίας	τούς λόγους τούτους	πεσῶν ἐξέψυξεν
D05	ἀκούσας	δὲ Ἀνανίας	τούς λόγους τούτους	παραχρῆμα πεσῶν ἐξέψυξεν
d5	audiens	autem Ananias	sermones hos	subito cum cecidisset obriguit
gig51	audiens	autem Ananias	hos sermones	cecidit et exspiravit
p54	audiens	autem Ananias	verba haec	momento cadens intra exspiravit

The Latin manuscripts are scrutinized in one variation place at a time in order to decide which Greek wording fits best at a given moment. In variation place A, as seen in table 1, the Latin *audiens* is paired with ἀκούων and *autem Ananias* with δὲ ὁ Ἀνανίας. The wording ἀκούσας of D05 is not found in any other manuscripts, and *audiens autem Ananias* is the standard beginning of this verse (as is ἀκούων δὲ ὁ Ἀνανίας in Greek witnesses); hence, we conclude that ἀκούων δὲ ὁ Ἀνανίας corresponds with *audiens autem Ananias*. In variation place C, all the Latin texts are judged to represent τούς λόγους τούτους, though there are variations in the Latin texts. Latin d5 and gig51 use the same words but in different order, whereas p54 has *verba haec*. However, the Latin *sermo* and *verbum* can both be used for λόγος. In variation place D, on the other hand, we differentiate gig51 from d5 and p54, so that *cecidit et exspiravit* of gig51 is translated as πεσῶν ἐξέψυξεν, following B03. The text in d5, *subito cum cecidisset obriguit*, is paired with *momento cadens intra exspiravit* of p54, both being taken as παραχρῆμα πεσῶν ἐξέψυξεν (“he at once fell down and died”), hence both agreeing with D05.⁴⁹

Granted, translating Latin texts into Greek is not an exact science as such, and another critic would end up making different decisions. For this reason, the critical editions of Acts and other available literature were extensively taken into account in this translating process, so that the decisions made here are based on a wider group of critics.

However, one cannot be confined to these two codices and other important witnesses agreeing with one or the other. There are two interesting lines of inquiry of the manuscript groups of Acts that deserve our attention. Both focus on the medieval minuscule manuscripts and deal with manuscripts of the secondary level of the B- and D-Clusters.

First, Barbara Aland and Georg Gäbel have studied a group of minuscule manuscripts, dubbed Group 614, which include manuscripts 383, 614, 1292, 1501, 1751, 2147, 2412, and 2652. According to their surveys, these minuscules form a textual basis for the main redaction (Hauptredaktion) of the D-Cluster, which explains the agreements between D05 and the witnesses in this group.⁵⁰ Epp, on the other hand, values two of these minuscules, namely, 383

⁴⁹ This is a common interpretation of this verse. See Rius-Camps and Read-Heimerdinger, “The Message of Acts,” 296; ECM 3/1.1: 136.

⁵⁰ According to the redaction hypothesis, the D-Text (Western) developed in stages and the main redaction was the result of conscious and deliberate recessional activity. Today, at least Gäbel has given up this hypothesis. See Barbara Aland, “Entstehung, Charakter und Herkunft des sogenannten westlichen Textes untersucht an der Apostelgeschichte,” *ETL* 62 (1985): 5–65, esp. 22–23, 63;

and 614, as secondary witnesses to the D-Cluster.⁵¹ Hence, Group 614 in its entirety is taken into account in this survey to see whether its elements form a cluster of their own or whether they cluster, for instance, with Do5.

Second, Thomas C. Geer has examined a group of twelve minuscule manuscripts called Family 1739: 206, 322, 323, 429, 453, 522, 630, 945, 1704, 1739, 1891, and 2200. Geer concluded that 945, 1704, 1739, and 1891 comprise the primary level of the textual family while 630 and 2200 make up the secondary level.⁵² These findings are highly intriguing to the present survey, since some of these manuscripts, 1739 and 1891, have been connected to the B-Cluster in Acts as the secondary witnesses to this textual group.⁵³ Hence, all of the minuscules (except 322)⁵⁴ of Family 1739 are taken into account here and the results compared with those of Geer.

What about the A-Cluster? Even though the scholarly discussion has been overly involved with the two older textual clusters of Acts, the A-cluster has plenty to offer regarding the later development, at least, of the text of Acts.⁵⁵ Hence, a selection was made of the manuscripts from the A-Cluster, with the chronological span being as wide as possible:⁵⁶ 014 (ninth century), 049 (ninth century), 0142 (tenth century), 35 (eleventh century), 69 (fifteenth century), 254 (fourteenth century), 876 (twelfth century), 1127 (twelfth century), and 1241 (twelfth century).⁵⁷ In addition, it was noted that particular minuscule manuscripts, rather unexpectedly, agreed with Do5 in certain verses in Acts 5;⁵⁸ therefore, these were also taken into consideration as a complementary set of manuscripts to those already mentioned: 044, 431, 913, 1611, 2138, 2344, 2401, and 2495.

Thus, altogether fifty-four manuscripts were chosen and summarized in table 2 below:⁵⁹

Georg Gäbel, “The Text of Φ^{127} (P.Oxy. 4968) and Its Relationship with the Text of Codex Bezae,” *NovT* 53 (2011): 107–52, esp. 150–52; “Western text,” 133–34.

⁵¹ Epp, “Textual Clusters,” 553–71; “Methodology for Isolating,” 235–39.

⁵² Thomas C. Geer, “Codex 1739 in Acts and Its Relationship to Manuscripts 945 and 1891,” *Biblica* 69 (1988): 27–46; *Family 1739 in Acts* (Atlanta: Society of Biblical Literature, 1994), 113.

⁵³ Alternatively, they have been connected to Category II (Egyptian text) by the Alands. Aland and Aland, *The Text of the New Testament*, 135, 335.

⁵⁴ The transcription of this minuscule was not available at the time of our analysis.

⁵⁵ Carroll D. Osburn and Gerald J. Dunker, for instance, conducted their studies using the threefold scheme to identify quotations of Epiphanius of Salamis and Athanasius of Salamis in Acts. See Carroll D. Osburn, *Text of the Apostolos in Epiphanius of Salamis*, NTGF 6 (Leiden: Brill, 2004), 38; Dunker, *Text of the Apostolos*, 43.

⁵⁶ Osburn and Dunker also included in the A-Cluster majuscules 020 and 025 with minuscules 1073 and 1352. These were not included here since 020 does not include Acts 5 and we did not have access to the other manuscripts or photos of them.

⁵⁷ These manuscripts belong to Category V (Byzantine), which is seen here as equivalent to the A-Cluster. See Aland and Aland, *Texts of the New Testament*, 110–38.

⁵⁸ See Pasi Hyttiäinen, “Evolving Gamaliel Tradition in Codex Bezae Cantabrigiensis, Acts 5:38–39: A Novel Application of Coherence-Based Genealogical Method (CBGM),” *TC* 24 (2019): 1–22. <http://jbtc.org/v24/TC-2019-Hyytiainen.pdf>.

⁵⁹ Four of these manuscripts contain only part of the chapter—namely, P45 (5:10–21, 30–39), 04 (5:35–42), 0189 (5:3–21), and h55 (5:23–42)—meaning that they are less reliable, an aspect accounted for in the analysis below. Since there are numerous Latin witnesses identified by the same number or letter, all Latin witnesses are indicated by a combination of the systems of Gryson and Nestle-Aland, resulting in the letter-number combination seen here. See ECM 3/2: 136.

Table 2. Manuscripts included in the phylogenetic analysis

B-Cluster	D-Cluster	A-Cluster	Group 614	Family 1739	Miscellaneous
P45	05	014	383	206	044
P74	d5	049	614	(307)60	431
01	gig51	0142	1292	323	913
02	p54	(33)	1501	429	1611
03	h55	35	1751	453	2138
04	08	69	2147	522	2344
0189	1884	254	2412	630	2401
1175		876	2652	945	2495
1409		1127		1704	
		1241		1739	
				1891	
				2200	

While more manuscripts could have been included, this study serves only as a preliminary survey of the applicability of the phylogenetic methods in Acts. Thus, the choice to include fifty-four witnesses is not meant to be comprehensive but to represent a sufficient case study. However, due to our decision to limit the survey to Acts 5, the results gained here are valid only for Acts 5 and not applicable to the text of Acts as a whole. As stated, some scribes had access to more than one exemplar, leading to contamination but also to block-mixtures, where a scribe copied a longer portion of the text using one exemplar but then, for some reason, changed to another exemplar. This means that the phylogenetic methods may relate particular manuscripts in Acts 5, but in the next chapter the connection cannot be detected due to a change of exemplar. An example of this is minuscule 33, which is called “Queen of the minuscules,” that is, containing text of great importance and quality.⁶¹ This minuscule has two faces in Acts, however. In 1:1–11:25 it is a good witness to the A-Cluster, but in 11:26–28:31 it is an excellent secondary witness to the B-Cluster (Egyptian).⁶² Hence, there was most probably a change of exemplars between Acts 11:25 and 11:26 in minuscule 33. Therefore, this minuscule is grouped here as a witness to the A-Cluster and not to the B-Cluster, as seen in table 2.

For these reasons, it must once again be underlined that this survey is not intended to comprise a comprehensive study of the textual history of Acts but rather methodological groundwork for further investigations.

Transcription and Collation of the Witnesses

After we have our data selected, the next step is to convert the manuscript texts into electronic form. This preliminary step is a relatively easy task, since many of the New Testament texts are already available via the New Testament Virtual Manuscript Room, edited by the Institut für Neutestamentliche Textforschung (INTF) in Münster.⁶³ This database does have one prominent weakness, however: it is dedicated to the Greek New Testament manuscripts. Since Latin manuscripts d5, gig51, p54, and h55 are also taken into account in this study, the Latin texts

⁶⁰ This minuscule was not included in Family 1739 by Geer but added here due to its high agreement rate with minuscule 453.

⁶¹ Aland and Aland, *Text of the New Testament*, 128; Aland and Wachtel, “Greek Minuscule,” 73.

⁶² Thomas C. Geer, “The Two Faces of Codex 33 in Acts,” *NovT* 31 (1989): 39–47, esp. 47.

⁶³ Of the 6464 witnesses catalogued by the INTF, only 2.61 percent have been transcribed. For detailed statistics, see <http://ntvmr.uni-muenster.de/>.

had to be translated and transcribed manually as a separate file. Additionally, since the Virtual Manuscript Room does not contain all the transcriptions needed here, we have transcribed minuscules 913 and 2401 to include them in the analysis.

After the transcribing process, the texts were collated—that is, aligned in the same file—so that verses with variant readings appear side by side. Previously this had to be done manually, but the process can now be done using the software CollateX, for example.⁶⁴ Thus, the painstaking task of collating the text, which had previously taken days, weeks or months, can now be done in a matter of minutes.

In the process of transcribing and collating the manuscripts, one particular question must be answered: what constitutes a variant reading? In other words, how should one define a difference between texts? CollateX identifies all variant readings, whether it include a typo, an absent *nu* from the end of a word, or the use of a different word altogether. Should all variant readings of different natures be treated the same? Are all variant readings significant? Here, we follow the common procedure of the phylogenetic analysis of manuscripts—that is, regularizing spelling and transcribing texts without punctuation, thereby omitting variants of a linguistic nature that are unlikely to contain relevant information about the history of the manuscript tradition.⁶⁵

Encoding the Texts

The texts are then encoded, enabling the phylogenetic software to process the data. Texts are often encoded with different letters or numbers to facilitate the identification of different changes within the manuscripts.⁶⁶ Scholars have also tested the idea of adding discrete weights to different variants, but this has had little effect on the results.⁶⁷ It should also be noted here that one of the chief criticisms of Lachmann's method has been its subjectivity—namely, in allowing critics themselves to select and evaluate the indicative errors.⁶⁸ If phylogenetic analysis should also require the critic to differentiate and assign weights to different types of variants, thus abandoning the intended principle of objectivity, this approach would start to come close to that of Lachmann's much-criticized method.

Hence, in this survey, all variants (except those of certain linguistic changes, as noted above) are treated equally—that is, all are counted equally. This was achieved by encoding the texts into sequences of numbers—for instance, in one place of variation, each reading is assigned a number: 0, 1, 2, et cetera. Using this method, all places of variation can be encoded, regardless

⁶⁴ CollateX is the successor to Peter Robinson's original software, Collate, available at <https://collatex.net/>. See Peter M. W. Robinson, "Collate: A Program for Interactive Collation of Large Textual Traditions, Version 1.1, Computer program distributed by the Computers and Manuscripts Project," in *Oxford University Computing Services* (1992).

⁶⁵ Spelling, for instance, differs significantly even within the most carefully written manuscripts, hence revealing little information about the relationships between manuscripts. Howe et al., "Manuscript Evolution," 123; James R. Royse, *Scribal Habits in Early Greek New Testament Papyri* (Leiden: Brill, 2008), 90.

⁶⁶ Linne R. Mooney et al., "Stemmatic analysis of Lydgate's 'Kings of England': A Test Case for the Application of Software Developed for Evolutionary Biology to Manuscript Stemmatics," *Revue d'histoire des textes, bulletin* 31 (2001): 275–97, esp. 287.

⁶⁷ This means that certain variants are given greater importance and weight. Spencer et al., "Weighting Kinds of Variants," 229, 238.

⁶⁸ William Robins, "Editing and Evolution," *Literature Compass* 4 (2007): 89–120, esp. 92.

of the number of variant readings.⁶⁹ The example from Acts 5:3 in table 3 demonstrates the procedure outlined. As can be seen, the texts are transcribed without punctuation, and the spelling is regularized (there is a missing *nu* at the end of *ειπεν* in minuscules 614 and 876).

Table 3. Encoding Acts 5:3

MS	No variation	Variation place	Encoding
B03	ειπεν δε	ο πετρος αναγια	0
614	ειπεν δε	πετρος αναγια	1
D05	ειπεν δε	πετρος προς αναγιαν	2
E08	ειπεν δε	προς αυτον ο πετρος αναγια	3
876	ειπεν δε	πετρος προς αυτον αναγια	4

In the beginning of this passage, there are no variations to be found. Hence, *ειπεν δε* is not encoded at all. It tells us nothing about how these witnesses are related, because they all agree. The agreement simply reveals that they are related, but this is something we already knew, since they all contain the Acts of the Apostles.⁷⁰ The other half of the passage, on the other hand, contains five different readings. The five readings were encoded as 0, 1, 2, 3, and 4 to differentiate them. Phylogenetic programs treat each different number as a different character state; that is, there are no hierarchies between them. Thus, all variants could be encoded with a single character-state change at the same place of variation. In case two manuscripts happen to include the same variant, they receive the same number. It is important not to encode each word separately, as this might add weight to a single change.⁷¹ The differing readings in table 3 are seen as a result of a single event, and thus we use a single number to also mark the multi-word changes in E08 and minuscule 876.⁷²

The forty-two verses of Acts 5 include altogether 279 cases where the manuscripts contain variant readings. In each instance of variant readings, every manuscript received a number referring to the variant it exhibits. This encoding resulted in a unique sequence of 279 characters (i.e., numbers) for each manuscript. However, before running the analysis, these sequences have to be converted to a NEXUS file, a format that is widely used in phylogenetics. The following example demonstrates the NEXUS file format and can be seen as instructions for the software to interpret the input data. The example in table 4 has five manuscripts, which contain altogether fifteen variant readings.

⁶⁹ Salemans based his computer-assisted stemmata on variation units comprising two variants, a method that is based on an underestimation of the capacity of present-day computers. Salemans, “Building Stemmas,” 22–35.

⁷⁰ See, e.g., Wassermann and Gurry, *New Approach*, 40.

⁷¹ See Howe, Connolly, and Windram, “Responding to Criticisms,” 61.

⁷² If, on the other hand, one suspects that beneath a multi-word change two or even three separate events have taken place, one should encode these variation places accordingly, using two or three different characters.

Table 4. Manuscript data converted to a Nexus file

```

#NEXUS
BEGIN TAXA;
  DIMENSIONS NTAX=5;
  TAXLABELS
      B03 D05 E08 614 876
;
END;
BEGIN CHARACTERS;
  DIMENSIONS NCHAR=15;
  FORMAT DATATYPE = STANDARD RESPECTCASE GAP = - MISSING = ? SYMBOLS = "0 1 2";
  MATRIX

      B03  000000000000012
      D05  010101010001001
      E08  101012000000100
      614  000000??1??1100
      876  000100100110000
;
END;

```

Using the NEXUS file, we specify the number of objects analyzed in this simple example (DIMENSIONS NTAX=5;), IDs of these objects (TAXLABELS B03, D05, E08, 614, 876), the length of the sequences (DIMENSIONS NCHAR=15), used symbols to differentiate variation (SYMBOLS="0, 1, 2"), et cetera. When this NEXUS file is downloaded to a phylogenetic program, the data is displayed as the following:

Table 5. Data matrix in a phylogenetic program

	1														
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5
B03	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
D05	0	1	0	1	0	1	0	1	0	0	0	1	0	0	1
E08	1	0	1	0	1	2	0	0	0	0	0	0	1	0	0
614	0	0	0	0	0	0	?	?	1	?	?	1	1	0	0
876	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0

In this simple example, we have five rows, each beginning with the ID of the manuscript, followed by the sequence of numbers. Each sequence is divided into 15 columns, which signify the places of variation in the manuscripts. In each column, different numbers represent the different readings found in the manuscripts. The program analyzes the data, column by column, while keeping track which manuscripts agree (contain the same character) in each column and finally, based on this data, it generates a phylogenetic tree (or network), using different methodologies. The whole dataset analyzed in this article contains 54 rows comprising 54 manuscripts, with each of the manuscripts represented by a 279-character sequence (using numbers when the texts are extant, a question mark when some portion of the text is missing, etc.), which transfer to 279 columns, which in turn represent the 279 variation places found between the manuscripts.

Phylogenetic Methodologies

Also necessary for this analysis is the selection of proper methods or algorithms, including a balanced understanding of the advantages and disadvantages of different methods, since phylogenetic trees are not self-generated but constructed in accordance with the rules or instructions determined by the algorithm applied.

An algorithm is a set of precise instructions or procedures to resolve a particular problem. Here, an algorithm has been designed to resolve (1) how the texts are related to one another based on the information inputted and (2) what is the most probable reconstruction of the evolution of a text from a single origin to the diverse state observed in the inputted texts.⁷³ In this study, we apply two methods for inferring phylogenetic trees, maximum parsimony (MP) and neighbor-joining (NJ), and one for constructing networks, NeighborNet. All these methods have been applied in textual criticism with promising results.⁷⁴

Maximum Parsimony (MP)

The maximization of parsimony is one of the oldest and most widely used phylogenetic methods in many fields—for instance, evolutionary biology, textual criticism, and cultural evolution.⁷⁵ The principle of MP states that the simplest explanation that explains the greatest number of observations should be preferred over more complex explanations. This principle is akin to Occam's Razor, which states that, all else being equal, the simplest hypothesis (in this case, the shortest tree) that explains the data should be selected. That is, the most accurate tree is most likely the shortest tree, which assumes the fewest amounts of mutations (evolutionary steps) or changes.⁷⁶

⁷³ Macé and Baret, "Why Phylogenetic Methods Work," 95.

⁷⁴ Robert J. O'Hara and Peter Robinson, "Report on the Textual Criticism Challenge 1991," *Bryn Mawr Classical Review* 3 (1992): 331–7; Caroline Macé, Philippe Baret and Peter Robinson, "Testing Methods on an Artificially Created Textual Tradition," in *The Evolution of Texts: Confronting Stemmatological and Genetical Methods, Proceedings of the International Workshop Held in Louvain-la-Neuve on September 1–2, 2004*, eds. Caroline Macé et al., *Linguistica Computazionale XXIV-XXV* (Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali 2005), 255–83; Catherine Eagleton and Matthew Spencer, "Copying and Conflation in Geoffrey Chaucer's Treatise on the Astrolabe: a Stemmatic Analysis Using Phylogenetic Software," *Studies in History and Philosophy of Science* 37 (2006): 237–68; Teemu Roos and Tuomas Heikkilä, "Evaluating Methods for Computer-Assisted Stemmatology Using Artificial Benchmark Data Sets," *Literary and Linguistic Computing* 24 (2009): 417–33; Teemu Roos, "Analysis of Textual Variation by Latent Tree Structures," *Data Mining (ICDM), 2011 IEEE 11th International Conference* (2011): 567–76.

⁷⁵ Walter M. Fitch, "Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology," *Systematic Zoology* 20 (1971): 406–16; Ruth Mace, Stephen Shennan and Clare J. Holden, eds., *The Evolution of Cultural Diversity: A Phylogenetic Approach* (Walnut Creek, CA: Routledge, 2010); Marko Halonen, "Computer-Assisted Stemmatology in Studying Paulus Juusten's 16th-Century Chronicle *Catalogus et ordinaria successio episcoporum Finlandensium*," *DSH* 31 (2016): 578–93.

⁷⁶ This does not, however, mean that most parsimonious approach is unable to describe complex evolutionary processes, such as the accidental rise of variants (homoplasy). In addition, the method does not assume that such phenomena are rare. It simply minimizes the number of homoplasies required to explain the genealogy of a given dataset. See, e.g., James S. Farris, "The Logical Basis of Phylogenetic Analysis," in *Advances in Cladistics II*, ed. Vicki A. Funk and Norman I. Platnick (New York: Columbia University Press, 1983), 7–36.

Underlying MP is the same principle outlined above: unnecessary assumptions are excluded by assigning equal weight to different types of variants—that is, all changes count as one change. These changes are summed up, and the tree with the lowest value is then designated as the most parsimonious tree.⁷⁷ This approach can sometimes result in trees of equal length—that is, trees with the same value—in which case the critic must evaluate and determine which to select. The other challenge faced when using MP is long-branch attraction (although other methods also have to deal with this problem), where objects are closely related not because they are related by descent but because they have undergone major, though unrelated, changes—that is, different lineages may have undergone the same changes by coincidence.⁷⁸ This is one error that must be identified during analysis.

Also, methods that have some optimality criterion, such as parsimony, should theoretically search out the optimal tree from all tree structures. Unfortunately, this is an all but impossible task due to the number of possible tree structures. Thus, different techniques are used for sampling trees.⁷⁹

Neighbor-Joining (NJ)

Neighbor-Joining is a distance-matrix method that infers phylogenetic trees from a matrix describing dissimilarities between each pair of objects. Pairwise distances are calculated from the matrix to construct a tree that minimizes the length of the branches and, subsequently, the length of the entire tree.⁸⁰

The starting point of NJ is a star-like tree that continues by searching for the least distant pair of witnesses—that is, the closest neighbors—and connecting them to one another by an ancestral node. The lengths of the branches that connect the witnesses to the newly generated ancestral node are defined by calculating the distance between one another and from this ancestral node to all remaining witnesses. This process continues until a true tree is obtained. The length of the generated branches depends on the distance value of each pair of neighbors, with the least distant values producing the shortest branch length and the most distant the longest branch length.⁸¹

The advantage of this method is its efficiency, making it possible to analyze vast data sets in a short amount of time. It has some disadvantages, however, as NJ finds the best solution that appears best at the moment, one step at a time, without accounting for the whole data set when searching for the optimal tree.⁸²

⁷⁷ Fitch, “Toward Defining,” 408–9, 413–14.

⁷⁸ Joseph Felsenstein, “Cases in Which Parsimony or Compatibility Methods Will Be Positively Misleading,” *Systematic Biology* 27 (1978): 401–10.

⁷⁹ We applied the stepwise addition method to find the optimal MP tree. Stepwise addition adds one sequence at a time to the tree until all are included and evaluates at each step all possible solutions, choosing the best option based only on the optimality criterion. In other words, it finds the solutions that seem best at any given moment. Ziheng Yang, *Molecular Evolution: A Statistical Approach* (Oxford: Oxford University Press, 2014), 83.

⁸⁰ This follows the principles of minimum evolution (ME), described in Kenneth K. Kidd and Laura A. Sgaramella-Zonta, “Phylogenetic Analysis: Concepts and Methods,” *American Journal of Human Genetics* 23 (1971): 235–52; Naruya Saitou and Masatoshi Nei, “The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees,” *Molecular Biology and Evolution* 4 (1987): 406–25.

⁸¹ Saitou and Nei, “Neighbor-Joining Method,” 407–9.

⁸² Yang, *Molecular Evolution*, 94.

NeighborNet (NN)

A phylogenetic tree implicitly assumes that once two lineages are created, they never interact with one another later on. A network, on the other hand, allows cycles, that is, paths that begin and end at the same node (as in fig. 1). In other words, a network can depict interactions between different lineages, as happens in the case of recombination or contamination.⁸³

There are several kinds of networks and methods to generate them, such as split decomposition, median networks and NeighborNet.⁸⁴ Even though all these methods have been applied to textual and other cultural data, NeighborNet seem to generate more resolved networks than, for instance, split decomposition; hence, we will here use the NeighborNet (NN) method.⁸⁵ Like the NJ method, NN is an agglomerative procedure, but it allows the graphs to contain cycles and thus be locally network-like. NeighborNet generates networks from a distance matrix of pairwise distances between the manuscripts. A split is a partition of the set of manuscripts into two subsets. When all possible splits are computed over a set of manuscripts, they can be either compatible or incompatible with one another. In the first case, there is a single way of connecting the manuscripts, which is a perfectly tree-like branching pattern. In the latter case, there are multiple ways of connecting the manuscripts, resulting locally in a network. NeighborNet summarizes the branching parts (edges) and the local networks (boxes) in a single graphical representation. Edges correspond to single unambiguous links while boxes embody conflicting connections.⁸⁶

Phylogenetic Analysis

Inferring Phylogenetic Trees and Bootstrapping

Both methods (MP and NJ) for inferring phylogenetic trees were run using the tree inference program PAUP* combined with bootstrap analysis, since this combination has produced reliable results in both phylogenetics and textual criticism.⁸⁷ All the tree inference techniques used here come with the package of PAUP* as defaults and need not be downloaded separately. PAUP* uses the NEXUS file format, as presented in table 4, for the input data.⁸⁸

Bootstrap analysis is a technique to test the statistical robustness of a tree by creating a

⁸³ Moulton and Huber, "Split Networks," 632.

⁸⁴ Hans-Jürgen Bandelt and Andreas W. M. Dress, "Split Decomposition: A New and Useful Approach to Phylogenetic Analysis of Distance Data," *Molecular Phylogenetics and Evolution* 1 (1992): 242–52; H. J. Bandelt et al., "Mitochondrial Portraits of Human Populations Using Median Networks," *Genetics* 141 (1995): 743–53; David Bryant and Vincent Moulton "Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks," *Molecular Biology and Evolution* 21 (2004): 255–65.

⁸⁵ Howe et al., "Manuscript Evolution," 147–52; Mahé Ben Hamed, "Neighbour-Nets Portray the Chinese Dialect Continuum and the Linguistic Legacy of China's Demic History," *Proceedings of the Royal Society B: Biological Sciences* 272 (2005): 1015–22; Eagleton and Spencer, "Copying and Conflation in Geoffrey Chaucer's Treatise on the Astrolabe," 237–68.

⁸⁶ Bryant and Moulton, "Neighbor-Net," 256–9.

⁸⁷ See, e.g., Masatoshi Nei, Sudhir Kumar and Kei Takahashi. "The Optimization Principle in Phylogenetic Analysis Tends to Give Incorrect Topologies When the Number of Nucleotides or Amino Acids Used Is Small," *Proceedings of the National Academy of Sciences – PNAS* 95 (1998): 12390–97, esp. 12396; Roos and Heikkilä, "Evaluating Methods," 422.

⁸⁸ David L. Swofford and J. Sullivan, "Phylogeny Inference Based on Parsimony and Other Methods Using PAUP*," in *The Phylogenetic Handbook*, 267–312.

number of additional trees based on samples of the original dataset. As mentioned, our dataset contains 54 rows and 279 columns. Now, the bootstrap algorithm included in PAUP* randomly selects columns from this dataset to generate a new set of 279 columns. Due to the random selection process, each column may be selected more than once or not at all. This process runs until the size of the original dataset (279 columns, each with values for 54 manuscripts) is reached; this is called a bootstrap sample, as seen in the example below with a smaller dataset:

Table 6. An example of bootstrapping

Original dataset										Bootstrap sample												
	1	2	3	4	5	6	7	8	9	10		6	2	7	9	1	8	5	3	4	1	
Bo3	0	0	0	0	0	0	0	0	0	0	Bo3	0	0	0	0	0	0	0	0	0	0	0
Do5	0	1	0	1	0	1	0	1	0	0	Do5	1	1	0	0	0	1	0	0	1	0	0
Eo8	1	0	1	0	1	2	0	0	0	0	Eo8	2	0	0	0	1	0	1	1	0	1	0
614	0	0	0	0	0	0	?	?	1	?	614	0	0	?	1	0	?	0	0	0	0	0
876	0	0	0	1	0	0	1	0	0	1	876	0	0	1	0	0	0	0	0	1	0	0

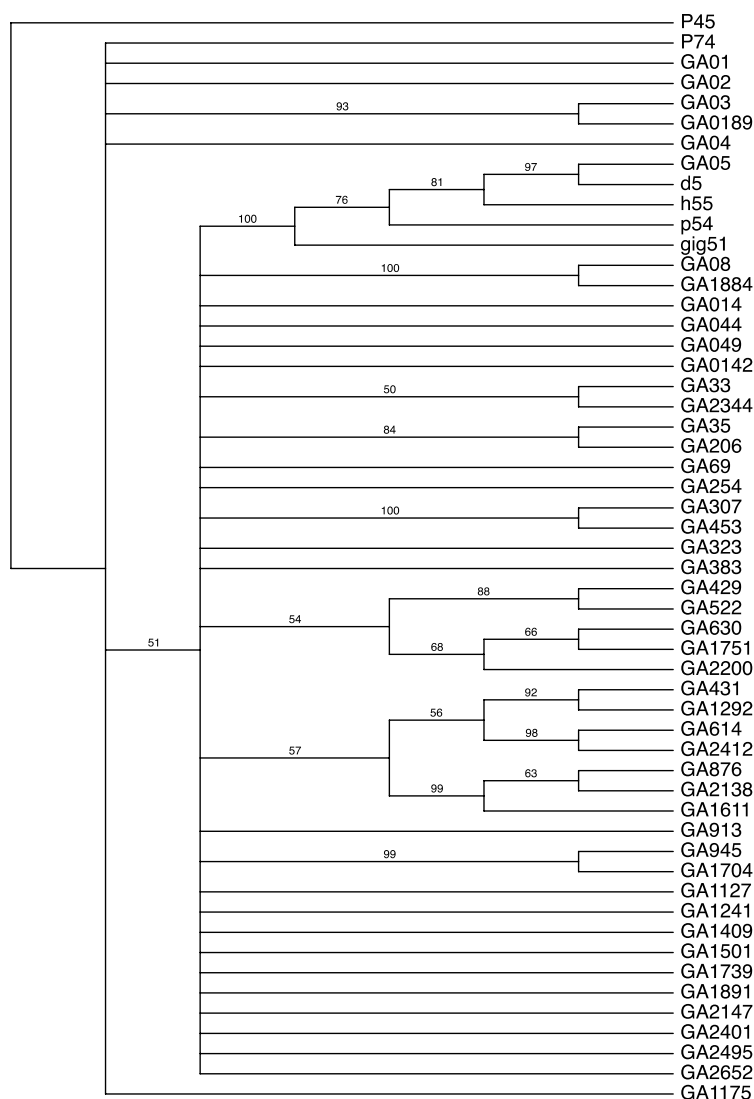
In this example, the algorithm randomly selected the columns, which led to the exclusion of column number 10 from the bootstrap sample, but at the same time column number 1 was selected twice.

After the size of the original dataset is reached, this process is repeated a number of times. In the following analysis, the process is repeated 1,000 times; that is, PAUP* generates 1,000 samples, and MP or NJ trees are constructed for every artificially generated bootstrap sample. In this process, the information within the columns about the relationships between the manuscripts remains untouched, but the resulting datasets differ from each other in regard to the types of variant readings selected into new sets of 279 columns. The random selection of the (types) of variant readings, repeated a number of times, aims to determine the probability of recurring splits in the tree structures.

Every split in the resulting tree is assigned a bootstrap value, computing the proportion of a given split occurring among the samples. If some split has a value of 65 percent, for instance, it means that this split is found in 65 percent of the trees generated by the bootstrap algorithm. When the value falls below 50, the algorithm groups the manuscripts together, signaling uncertainty about their relationships. While it is not an easy task to interpret the generated values, branches or groups that are given above 75 percent support can be treated with considerable confidence.⁸⁹

⁸⁹ In addition, 1000 samples seems to be a sound number for our analysis. See Yves Van de Peer, "Phylogenetic Inference based on Distance Methods," in *Phylogenetic Handbook*, 142–60, esp. 157–59; Swofford and Sullivan, "Phylogeny Inference based on Parsimony," 297; Yang, *Molecular Evolution*, 172; concerning the question why bootstrapping works, see Joseph Felsenstein, "Confidence Limits on Phylogenies: An Approach Using the Bootstrap," *Evolution* 39 (1985): 783–91.

Figure 2. Tree of Acts 5, Maximum Parsimony



In figure 2, representing the MP tree, the witnesses are located at the tips of the branches, identified using the Gregory-Aland numbering system; either P or GA⁹⁰ is in front of the manuscript number for Greek witnesses, the remainder being Latin manuscripts. The tree is an unrooted one, generated by PAUP*⁹¹; that is, it does not have a starting point and must not be confused with a stemma, which always has, in the traditional sense of the term, an archetype or a root.⁹¹

The tree inference methods used here usually generate strictly bifurcating trees, that is, unrooted binary trees with exactly three neighbors at each internal node. This is the standard way to represent evolutionary relationships between organisms.⁹² The MP tree above, however, reveals that several branches lack a bootstrap value, meaning that they are simply grouped together since there is little statistical support for the bifurcations, the bootstrap value being under 50 percent.

⁹⁰ In the following discussion, we use the numbers only, without “GA,” to identify Greek manuscripts.

⁹¹ West, *Textual Criticism*, 32.

⁹² Vandamme, “Basic Concepts,” 28.

Hence, this is a multifurcating tree, which is not a bad thing as such since it is well known that scribes often made several copies from a single exemplar. A multifurcating tree is able to depict these events, since it allows more than two descendants per given ancestral node.⁹³ However, the amount of multifurcations and the low support values (51, 54, 57, etc.) in the splits prevent us from making justifiable observations regarding the textual tradition of Acts 5 as a whole, which may signify that there are contradictory elements in the data.⁹⁴

This is not a surprising outcome as such, since it is acknowledged that the textual tradition of the New Testament in general and the Acts of the Apostles in particular is highly contaminated; that is, every manuscript is a mixture of different textual states or influenced by different textual clusters.⁹⁵ As stated earlier, scribes used the manuscripts that were at their disposal, with greater or lesser care, whether they contained older or younger text.⁹⁶ In addition, manuscripts of Acts 5 contain (as does every New Testament text) several accidental agreements; that is, scribes made the same change independently from one another.⁹⁷

Nevertheless, we can make some observations about particular parts of the tree. It is worth noting that the possible contradictory elements seem to be present even within the oldest witnesses, that is, the primary witnesses to the B-Cluster, P45, 01, 02, and 03. Only 03 and 0189 are connected with a relatively good support value of 93. The primary witnesses to the B-Cluster are separated from the rest of the tree, but their mutual relationships cannot be substantiated. On the other hand, the primary witnesses to the D-Cluster—gig51, p54, d5, h55, and 05—are grouped together with relatively high bootstrap values, all splits exceeding 75 percent. In fact, this is the only textual cluster from the traditional groups that clearly emerge from the tree and are supported by bootstrap values having statistical significance. This is somewhat surprising, considering that critics have in recent discussions argued to abandon the D-Cluster or the D-text as a distinguishable textual entity.⁹⁸ Here it is the most distinguishable textual cluster, depicting the Latin gig51, p54, d5, and h55 and Greek 05 as closely related manuscripts. The secondary D-Cluster witnesses, 08 and 1884, are not located on the same branch but nearby nevertheless.

The other branches have partially lower values, one containing minuscules 431, 614, 876, 1611 with 2138, 1292, and 2412. To find these minuscules grouped together is not surprising, for they are also found together in the diagrams of CBGM in Acts 5.⁹⁹ We also find some of Family 1739 grouped together. In one branch, we have the primary witnesses 945 and 1704 and as secondary witnesses minuscules 630, 1751 (not grouped with Family 1739 in previous studies) and 2200, including a kind of third-rank witnesses, that is, 429 and 522. On the third branch, we find minuscules 307 and 453 with 100 percent support value.¹⁰⁰

As pointed out, there are many uncertainties in the MP tree with little statistical support

⁹³ Howe, Connolly and Windram, “Responding to Criticisms,” 63.

⁹⁴ Vandamme, “Basic Concepts,” 29; Yang, *Molecular Evolution*, 175.

⁹⁵ The used terminologies partly depend on the theoretical framework from which one is approaching the textual tradition. See Gerd Mink, “Problems of a Highly Contaminated Tradition: The New Testament Stemmata of Variants as a Source of a Genealogy for Witnesses,” in *Studies in Stemmatology II*, 13–85, esp. 13; “CBGM Presentation,” 29–31; Epp, “Methodology for Isolating,” 228–32.

⁹⁶ Aland and Wachtel, “Greek Minuscules,” 72.

⁹⁷ Concerning this phenomenon, see, e.g., Wasserman and Gurry, *New Approach*, 25–26.

⁹⁸ Parker and Pickering, “4968 Acta Apostolorum,” 6; Gäbel, “Western Text,” 133–34; Wachtel, “On the Relationship,” 147.

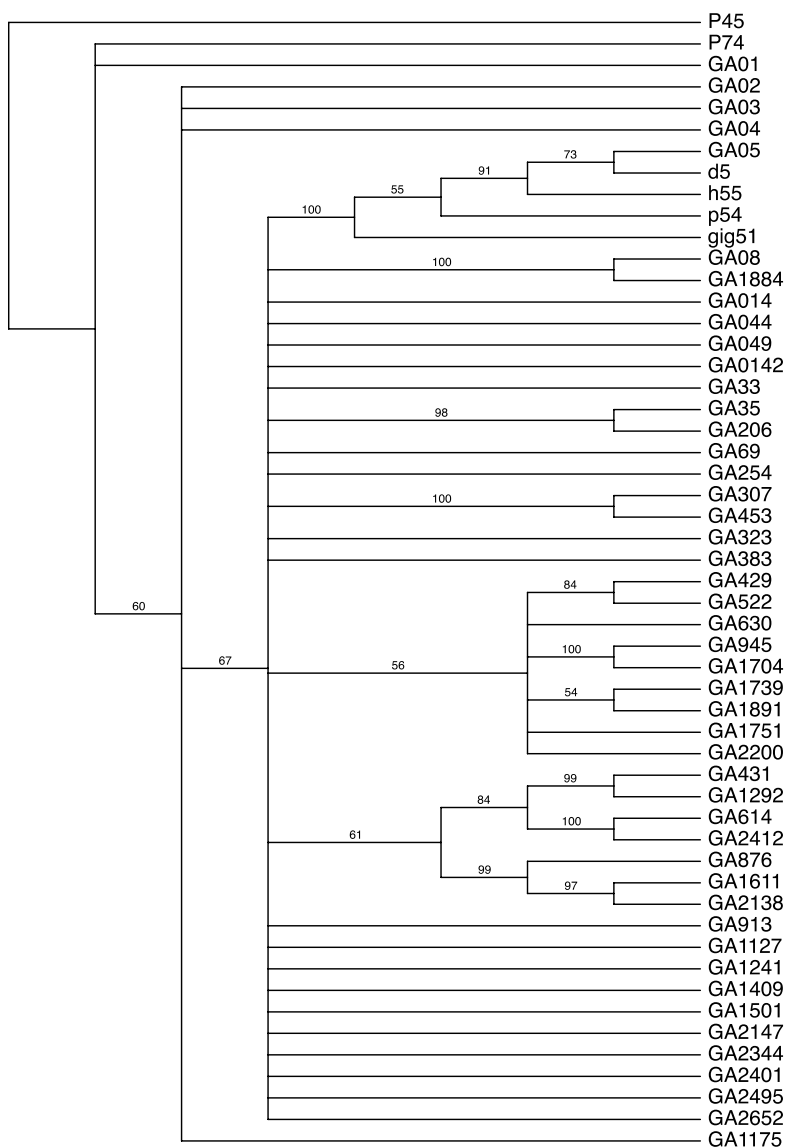
⁹⁹ This diagram is found under the module “Coherence and Textual Flow,” <https://ntg.uni-muenster.de/acts/ph4/coherence/1>.

¹⁰⁰ See Geer, *Family 1739*, 113.

to make further deductions based on the data we have at our disposal. It is therefore pointless to push the analysis further, since the evidence is simply not there. Hence, we now move to the neighbor-joining tree, which is, unlike the MP tree, constructed using a distance matrix of pairwise distances between manuscripts.

Once again, this is an unrooted tree¹⁰¹ and must not be confused with a stemma or treated as one. There are once again several branches without bootstrapping values. Having said that, this tree is more resolved; that is, there are fewer multifurcating branches. Even though there are more bifurcations detected—that is, splits in two branches—their bootstrapping values are low: 56, 60, 61, 67, et cetera.

Figure 3. Tree of Acts 5, Neighbor-Joining



¹⁰¹ The NJ tree only contains fifty-three manuscripts, for it was necessary to exclude 0189 from the analysis. Majuscule 0189 contains 5:3–21 and the Latin h55 5:23–42. Hence, there are no common texts between these witnesses. The NJ method constructs trees from a distance matrix containing distances between pairs of manuscripts. If there are no common characters preserved, PAUP* assigns the value zero as a default, meaning that in every run, 0189 and h55 were connected as closest relatives, even though they contain different portions of Acts 5. We opted to delete 0189 due to the relative importance of h55 to the D-Cluster.

The witnesses to the B-Cluster—P74, 01 and 02, 03 and 04, for instance—are grouped, but their exact relation remains unresolved once again. Also, the D-Cluster is separated from the rest of the tree. What is noteworthy is that the split leading to the cluster has here, as in the MP tree, 100 percent support value, meaning that this split is present in every single tree, that is, in 1000 artificially generated trees. This means that, at least statistically, the witnesses to the D-Cluster have a common lineage (i.e., they are closely related to one another).

One must ask, however, whether we have here an example of long-branch attraction. As stated above, in this phenomenon objects are closely related not because they are related by descent but because they have undergone major, though unrelated, changes—that is, different lineages may have undergone the same changes by coincidence. Most of the common variants detected in the D-Cluster manuscripts concern multiple words, for example, meaning that a genealogical relationship must exist between them—that is, they likely did not undergo the same changes by accident. An example of this is found in 5:39, where 05, d5, and h55 contain the reading “neither you, nor kings, nor tyrants; thus, keep away from these men.”¹⁰² Different scribes did not invent this long reading independently but adopted it from an exemplar. Thus, the D-Cluster branch was most likely not generated by long-branch attraction but because these manuscripts are genuinely related to one another.

The other feature of great interest is the Family 1739 cluster in the NJ tree. Now we have all the primary and secondary witnesses to Family 1739 grouped together, with 429 and 522. Again, we find minuscule 1751 grouped with this cluster, but the exact relation between it and the other manuscripts in this cluster remain indecisive. Previously, this minuscule was grouped with Group 614. Three of the manuscripts included in Group 614—minuscules 614, 1292, and 2412—are found in the same group with the same minuscules as in the MP tree with good statistical support (except for the split leading to this cluster).

However, the amount of multifurcations and low support values raise the question whether there indeed are contradictory elements present in the data. If so, one has to ask whether the complexity of the tradition can be fitted into a tree. In the face of recombination (or contamination), as suspected here, the data may support two or more contradictory tree structures. In such cases, the tradition might not always be best represented by a bifurcating or multifurcating tree but rather a network.¹⁰³

Constructing Networks Using NeighborNet

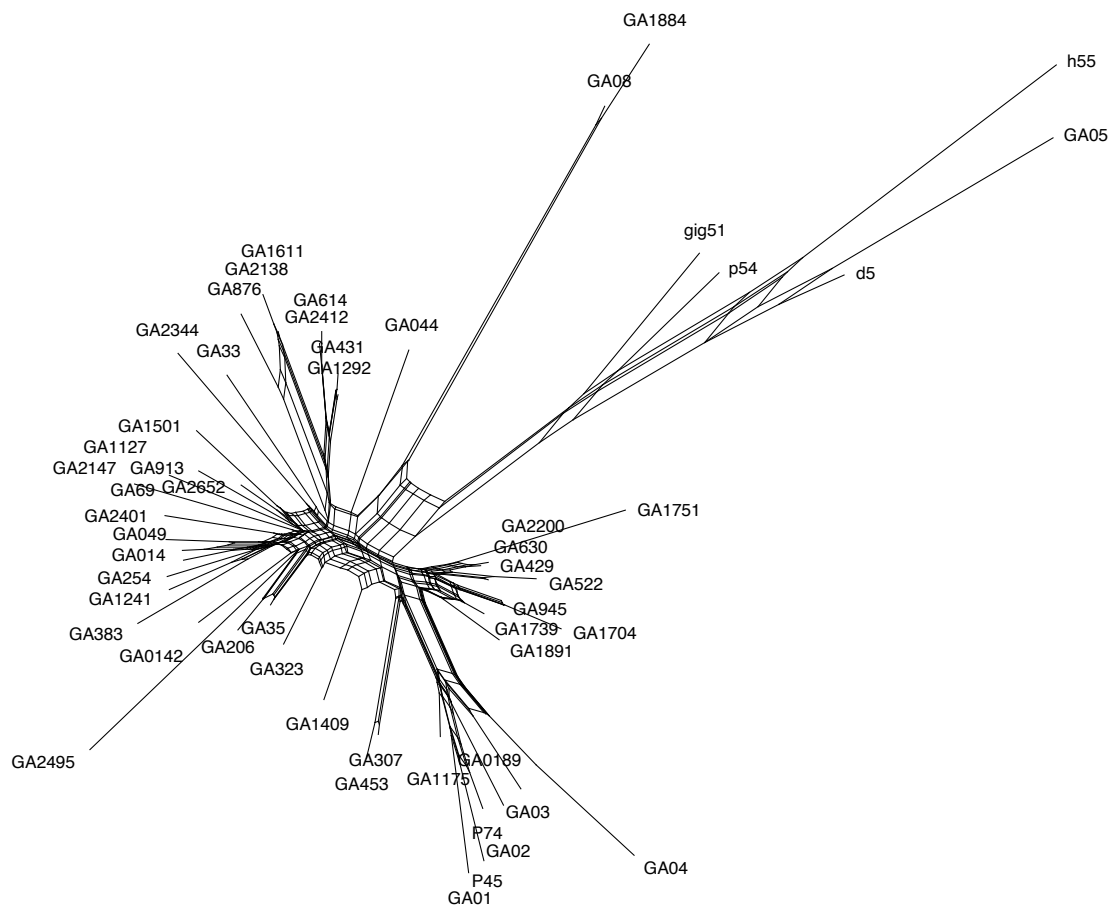
The NeighborNet method was run using SplitsTree phylogenetic software.¹⁰⁴ Now, what we have in figure 4 is not a tree—that is, straight parallel edges connecting manuscripts—but a network where several boxes are located between the witnesses. Every one of these boxes represents conflicting signals, meaning that witnesses can be connected with one another in a variety of ways. The lengths of edges are proportional to the amount of change occurring along them; that is, more changes lead to longer edges.

¹⁰² Hyytiäinen, “Evolving Gamaliel Tradition in Codex Bezae,” 12–17.

¹⁰³ Moulton and Huber, “Split Networks,” 631–32.

¹⁰⁴ Daniel H. Huson and David Bryant, “Application of Phylogenetic Networks in Evolutionary Studies,” *Molecular Biology and Evolution* 23 (2006): 254–67.

Figure 4. Network of Acts 5, NeighborNet



Because we suspect that there might be contamination responsible for the conflicting signals, when generating the boxes, we apply recombination detection methods to our dataset. The point of using these tools is to find out whether there are signs of recombination and whether they can be isolated to a particular region within the sequences and then excluded from the analysis. There are a number of recombination detection applications, but here we used TOPALi, for it contains several detection methods that can be used to crosscheck the findings.¹⁰⁵ The methods of detecting recombination events used here are PDM (Probabilistic Divergence Measures), LRT (Likelihood Ratio Test), and DSS (Difference of Sums of Squares).¹⁰⁶

These methods use a sliding window approach in which a fixed-size window is moved along the alignments, and in each position of the window a tree is estimated, using the Markov chain Monte Carlo approach (PDM), neighbor-joining (LRT), or by minimizing the sum of squares (DSS). While the window slides along the sequences (and in each position a tree is estimated), the differences between each position of the window are calculated and recorded in a graph (see fig. 6).¹⁰⁷ These methods are exceedingly computationally demanding; thus, we cannot take all the sequences into the analysis but rather a representative sample. We use

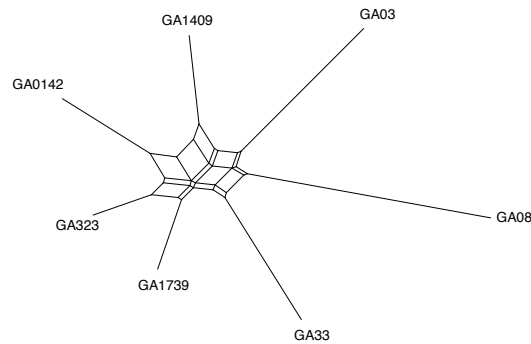
¹⁰⁵ I. Milne et al., "TOPALi: Software for Automatic Identification of Recombinant Sequences within DNA Multiple Alignments," *Bioinformatics* 20 (2004): 1806–7.

¹⁰⁶ More information about these methods, see D. Husmeier and F. Wright, "Probabilistic Divergence Measures for Detecting Interspecies Recombination," *Bioinformatics* 17, Suppl. 1 (2001): S123–31; G. McGuire and F. Wright, "TOPAL 2.0: Improved detection of mosaic sequences within multiple alignments," *Bioinformatics* 16 (2000): 130–34.

¹⁰⁷ McGuire and Wright, "TOPAL 2.0," 130–4.

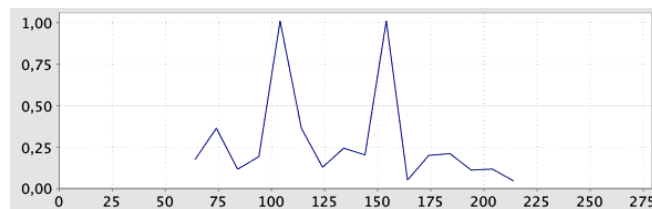
NeighborNet to select a sample of seven manuscripts, which are well distributed across the network and may constitute a useful subset to investigate evidence for recombination.¹⁰⁸ The selected manuscripts are 03, 08, 0142, 33, 323, 1409, and 1739. We can see in figure 4 that these manuscripts are spread across the network; that is, they have a good distribution. The network generated by NeighborNet of the selected manuscripts is the following:

Figure 5. NeighborNet of Acts 5 (7 witnesses)

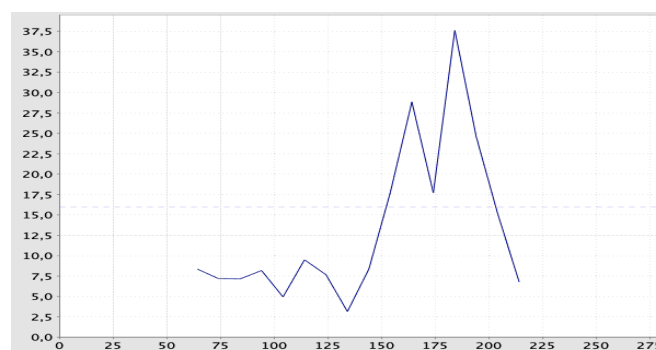


The recombination detection analysis was done using these seven manuscripts only and summarized in figure 6. The numbers running below the graphs correspond to the columns in the sequences, that is, running between 0–279. Each spike in the graph signals changes in the values detected, thus suggesting recombination events. These tools are all unanimous in that there are indeed recombination events detected in our dataset and multiple events in fact.

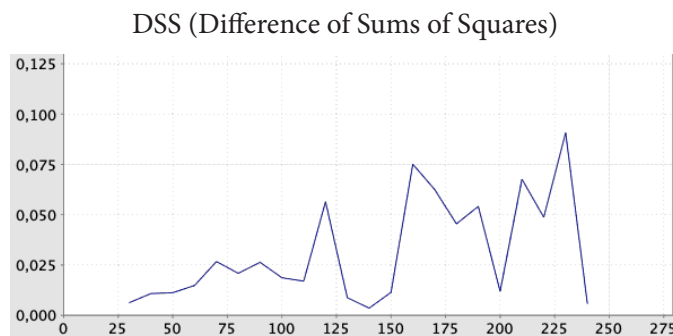
Figure 6. Results of the recombination detection analysis
PDM (Probabilistic Divergence Measures)



LRT (Likelihood Ratio Test)



¹⁰⁸ Moulton and Huber, “Split Networks,” 648.

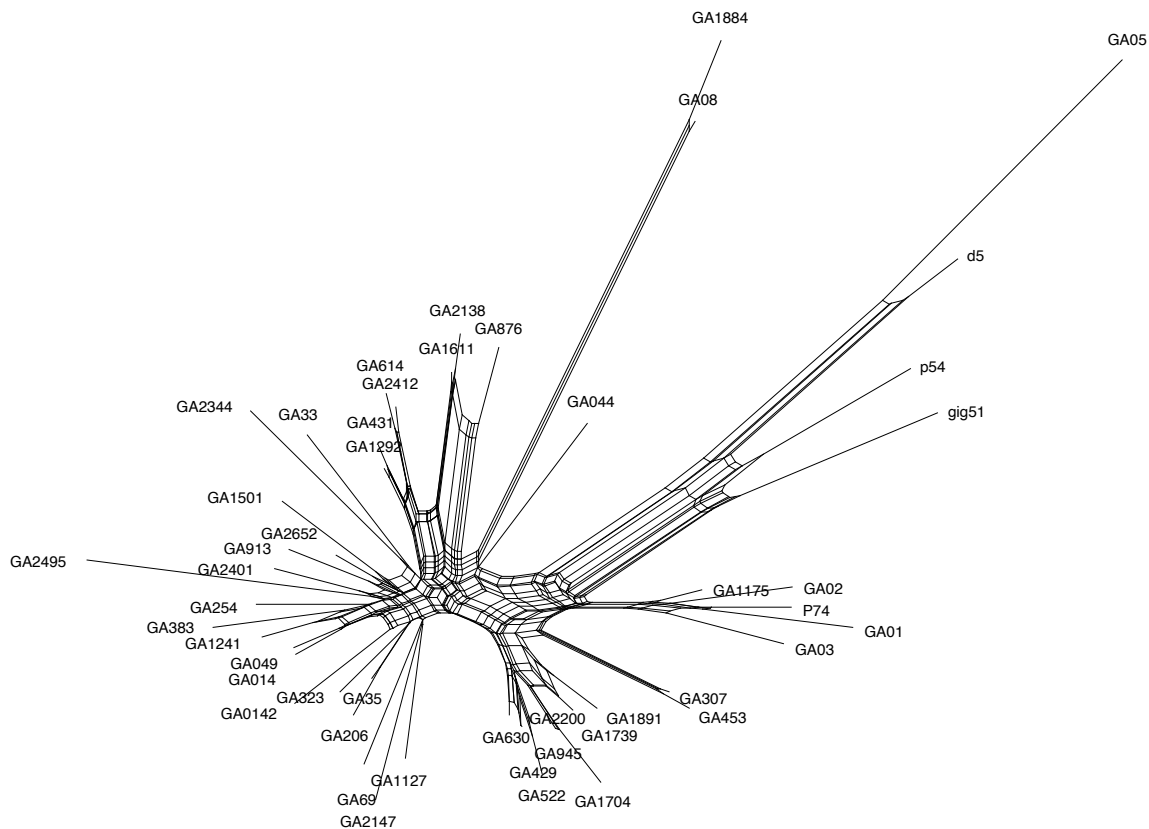


Also, it seems that the recombination events cannot be isolated to a particular region in the sequences, but rather they are present everywhere. Different window sizes were tested, ranging from 127 (the smallest possible window size allowed by the PDM method) to 100 (LRT) and down to 50 (DSS), in order to specify more carefully where the recombination events take place, resulting in the graphs above. For instance, taking the columns (or sites) where the clearest spikes are in the first graph (PDM), sites 85–115 and 140–160, and excluding them from the analysis, there is no change in the NeighborNet graph in figure 5. The same holds true if we exclude the entire block of sites 130–210 as suggested by the graph of the LRT method. In other words, according to the used methods, our dataset is heavily contaminated and there is no way to filter all the sites where recombination is detected, as that would mean excluding the majority of the columns and call into question the validity of the entire analysis. Hence, we must return to the NeighborNet graph in figure 4 and take a closer look at the manuscript groupings.

We can refine the graph a bit by excluding all the manuscripts from the graph that are heavily lacunose (P45, 04, 0189 and h55), for it seems that they cause ambiguity, that is, boxes around them. The same holds true in the case of two particular witnesses, 1409 and 1751, but here we seem to have a case of exemplar change. This change can be identified with relative accuracy using NeighborNet, for if we exclude columns 0–99, nearly all the boxes are deleted in the case of 1751, and it is placed right in the middle of Family 1739. On the other hand, by excluding the first 100 columns, the boxes are deleted in front of 1409, and it is placed between 307 and 1175. If we include the first 100 columns in the analysis only, 1409 is related to 69 and 1751 to a group of three manuscripts: 876, 1611, and 2138. The first 100 columns comprise Acts 5:1–16; thus, the exemplar change seems to have occurred between 5:16 and 5:17.¹⁰⁹ However, because we are focusing here on Acts 5 as a whole, all six witnesses (P45, 04, 0189, h55, 1409, and 1751) are excluded from the final NeighborNet graph below.

¹⁰⁹ In fact, NeighborNet has proven to be a useful tool for detecting the exemplar change. See Eagleton and Spencer, “Copying and Conflation in Geoffrey Chaucer’s *Treatise*,” 253–59.

Figure 7. NeighborNet of Acts 5 (48 witnesses)



The exclusion of these manuscripts did not delete the boxes, but now we can discern the groupings and the general structure more clearly, enabling us to make some observations. By excluding fragmentary witnesses to the B-Cluster, P45, 04, and 0189, this cluster seems much more tightly grouped with almost none of the boxes left. Encouraged by this, it was tested whether running the MP and NJ methods with only forty-eight witnesses would result to more resolved trees. Unfortunately, this did not have any effect on the tree structure. The MP and NJ methods were unable to reconstruct, for instance, the relationships between the witnesses to the B-Cluster, but they were once again simply grouped together. The exclusion of h55 did not change much of the grouping of the D-Cluster.

The majority of the witnesses to Family 1739 are again clustered together with all the primary and secondary witnesses found in this cluster, supplemented by the third level of the family, that is, manuscripts 429 and 522.¹¹⁰ This seems to confirm the findings of Geer that these manuscripts indeed constitute a textual family or cluster of their own. This is noteworthy since, according to Geer, chapters 1–6 are the least affected by the Family 1739 readings in Acts. In spite of this, NeighborNet (and NJ) grouped these manuscripts together. Also, as Geer suspected, minuscules 206 and 323 have stronger affinities to the A-Cluster in this section of Acts, which are separated from the core members of Family 1739 in figure 7.¹¹¹

On the other hand, we seem to have two parts or sides in the network. The one on the right comprises three clusters: D-Cluster, B-Cluster, and Cluster 1739 (as it will be called henceforth). On the left side, there are two larger groupings, the one between 33 and 1127, and the other, between 1292 and 876. The manuscripts 08, 044, and 1884 fall somewhat apart from

¹¹⁰ The same grouping can be also found in the diagrams of CBGM. See <https://ntg.uni-muenster.de/acts/ph4/coherence/1>

¹¹¹ Geer, *Family 1739*, 74, 83, 87, 113.

these two groups. It is underlined that there is no direction in this network—that is, there is no starting point—and it must not be confused with a rooted tree or stemma in this respect.¹¹² However, one can theorize that the older part of this network is on the right side, where the B- and D-Clusters are located. The network indicates that the B-Cluster and Cluster 1739 are separated; thus, we are dealing here with two separate textual groups. Granted, all of the witnesses to Cluster 1739 have a high level of agreement, exceeding 90 percent in Acts 5.¹¹³ However, these witnesses also often agree with the members of the B-Cluster, particularly with 1175. Hence, we may be on firmer ground by seeing Cluster 1739 as a subgroup of the B-Cluster.¹¹⁴

The witnesses between 33 and 1127 can be interpreted as the A-Cluster. All but one (876) of the witnesses classified as belonging to the A-Cluster before the analysis are located in this larger cluster, supplemented by the majority of the witnesses to Group 614. This confirms Geer's suggestion that in this section of Acts minuscule 33 has strong affinity to the A-Cluster.¹¹⁵ Furthermore, the members of Group 614 seem to be primarily witnesses to the A-Cluster and do not form their own group. On the other hand, the group between 1292 and 876 was identified by means of every method applied here, and hence we may have here a subgroup of the A-Cluster.

Conclusions

What we learned from this survey? The phylogenetic methods indeed yield relevant information about the textual tradition of Acts 5. They confirm what was already known: contamination is widespread and a difficult problem to deal with. One single chapter of Acts challenged the phylogenetic methods to such an extent that the data could not be fitted to a single tree. The relations between manuscripts are so complex that it was impossible to substantiate the exact relations, for instance, between the witnesses to the B-Cluster, which is, after all, considered to contain very ancient text. The idea that the data cannot be depicted in a single tree, whether a bifurcating or multifurcating tree, is a factor that needs to be taken seriously, since it means that the development or the evolution of the textual tradition of Acts 5 is not tree-like. This, in turn, casts doubts on stemmatological studies conducted in New Testament textual criticism, for in these studies textual development has been primarily depicted using tree-like structures.

This is where networks come into play. Networks can depict complex structures without forcing the data into a tree structure. Here lie the greatest prospects, we think, in applying the phylogenetic approach to Acts or the New Testament in general. As demonstrated, networks—NeighborNet in particular—allow the textual critic to discern ambiguities in the manuscript transmission, detect exemplar changes and see clustering patterns within the textual tradition. In our case, NeighborNet was able to discern several clusters that corresponded rather well with previous studies, except in the case of Group 614. NeighborNet found the two ancient textual groups, B- and D-Clusters, supplemented with Cluster 1739. If Cluster 1739 is interpreted as a subgroup of the B-Cluster, then we have here the two early textual forms. NeighborNet also aggregated a large group of later witnesses, which can be labeled as the A-Cluster.

¹¹² Eagleton and Spencer refer to NeighborNet as a stemma, which is somewhat misleading, in "Copying and Conflation in Geoffrey Chaucer's Treatise," 245.

¹¹³ This again agrees with the data of CBGM. See n. 99.

¹¹⁴ The witnesses to Cluster 1739 are often characterized as a secondary Alexandrian or Egyptian. See Geer, *Family 1739*, 113; Osburn, *Text of the Apostolos in Epiphanius*, 38; Donker, *Text of the Apostolos in Athanasius*, 43.

¹¹⁵ Geer, "Two Faces of Codex 33," 47.

We can ask, does the evidence gathered here and the methodology used give us grounds to abandon the text-type theory in Acts 5 or the two-text concept? The answer is no. The phylogenetic methods discerned the two ancient clusters, regardless of how we interpret the existence of Cluster 1739. Furthermore, the phylogenetic approach gives us reason to propose that the only way to tackle contamination is by using networks and clustering. This is due to the realization that the tree inference methods MP and NJ were unable to relate the majority of individual manuscripts to one another in Acts 5.

As stated, this was a preliminary analysis, offering methodological groundwork for further studies. The next step is to apply the network approach found to be the most promising method here to several chapters of Acts to determine whether our findings apply elsewhere in Acts.