

Luigi Bambaci

IS A STEMMA POSSIBLE FOR THE HEBREW BIBLE?
TOWARDS A GENEALOGY OF MEDIEVAL MANUSCRIPTS
THROUGH PHYLOGENETIC ANALYSIS

Synopsis

We propose here a computer-assisted approach to stemmatic analysis of the Hebrew Bible, employing a quantitative method common in many philological domains, but to our knowledge unprecedented in Old Testament studies. To be specific, it is our intention to explore the potential of phylogenetic methods, taking as our case study Kennicott's collation of Qohelet. The need to systematise the medieval documentation of Qohelet has arisen in the preliminary stages of our work on producing a born-digital critical edition of that book.

The article is organized as follows: After a general introduction (§ 1), we pause to summarize the state of affairs today in the field of stemmatic analysis of medieval biblical Hebrew manuscripts (§ 2). We then present the data we have gathered (§ 3) and the method we used to analyse them (§ 4). Finally, we describe and discuss the results obtained (§§ 5,6), concluding with a few speculations as to future research (§ 7).

1 *Introduction*

The aim of our analysis is to define the physiognomy of the medieval tradition of Qohelet (Q) and in particular to propose a reconstruction of witness relationships through the definition of textual groups or families. The task we

set ourselves pertains therefore to stemmatology, a branch of textual criticism traditionally little frequented by scholars of the Hebrew Bible (HB), and considered by the vast majority to be inapplicable to the case of the medieval tradition.

The source from which we have derived our data is the 18th-century Kennicott collation (KN).¹ The choice of KN – or alternatively, of De Rossi (DR)² – is even today an obligatory choice which imposes severe limitations, above all the exclusion of vocalization and punctuation, as well as of the Massora and of manuscripts (MS/MSS) not known to the collators at the time. An altogether new collation of the medieval documentation designed to fill these gaps would have far exceeded the resources allocable to this study – not to mention the competence of this writer. Moreover, it would not have been able to guarantee the extensiveness required by our quantitative analysis and provided precisely by the aforementioned collations.

Of the two, we have chosen to work with KN because it is more comprehensive. As is known, DR cites only variants that seem to him the most important and above all those that also occur in the ancient Versions.³ This makes it impossible to check his selection criteria and therefore to ascertain the extent to which the sample chosen by him is actually representative of the surviving textual tradition. Furthermore, unlike KN, DR does not clearly specify which witnesses were collated in full and which only sporadically.

The citations from the secondary literature have all been translated by us. Some particularly frequent names or entities are indicated by acronyms from their second occurrence in the text onward. The acronyms 'KN' (Kennicott) and 'DR' (De Rossi) indicate metonymically either the person or the work.

¹ B. KENNICOTT, *Vetus Testamentum Hebraicum cum variis lectionibus*, I-II, Oxford 1776-1778. Qo-

helet is in vol. II, pp. 549-61.

² G.B. DE ROSSI, *Variae lectiones Veteris Testamenti*, I-IV, Parma 1784-1788 (Qohelet is in vol. III, pp. 247-64, and IV, *Appendix*, pp. 237-8); ID., *Scholia critica in V.T. libros, seu supplementa ad varias sacri textus lectiones*, Parma 1798, pp. 130-2.

³ See DE ROSSI, *Variae lectiones*, I, pp. xviii, xliii-xliv.

Although more suitable as a sample for a stemmatic investigation, even KN is not without its problems. These concern in general the accuracy with which variants were recorded in the critical apparatus and in particular the lack of distinction between first- and second-hand variants.

To obviate these difficulties, at least in part, we have chosen to re-collate several of the MSS examined by KN. Finally, we have collated a certain number of altogether new MSS, chiefly Oriental, in order to broaden the representativeness of the sample.

This re-examination of the Kennicottian MSS has proven extremely useful because it has allowed us to understand better the *modus operandi* of the collators and to correct errors and inaccuracies. This clearly does not alter the fact either that errors may persist or that others may have been introduced during both the process of revision and the collation of the new material. Nevertheless, the quantitative nature of the method we decided to employ should be able to guarantee a certain reliability of the final results, canceling out the possible – and hopefully sporadic – presence of noise.

A final and important limitation concerns the absence of data from the ancient Versions, which are a prerequisite to understanding whether the variants in the Hebrew codices are pre-Masoretic, and, if so, original or not. To have such data at our disposal, an exhaustive collation of the tradition of Q as a whole would be required, something which none of the editions of the HB yet allow for.

The method we have chosen is based on the genealogical model and is borrowed from biology. It is essentially quantitative and has been implemented thanks to the use of the computer. Although abstract, it tries to take into account both the logic of the copying process and the peculiarities of the HB medieval tradition.

Despite its limitations, we contend that this method forms a valid tool for stemmatic analysis, at a minimum as a first systematization of the material and as a phase preparatory to a qualitative analysis.

Anticipating some of our results, we may immediately declare that our method has indeed permitted the separation of witnesses into discrete groups, albeit at various levels of confidence. Some of these groups are questionable because they are based on variants with weak kinship-revealing power; others seem more certain, both on the basis of the number of shared variants and on the basis of type – in some cases, even on the basis of characteristic variants – and call for further study. The general picture that emerges from such groupings is well in line with prior studies on textual history of the HB. The distinction between witnesses of the *textus receptus* (TR) type – mostly Sephardi and Oriental – and those of the *anti-receptus* (AR) type – mostly Italian and Ashkenazi – is also applicable to Q.

2 *Status quaestionis*

Following the publication of KN and DR, a heated debate ensued as to the usefulness of the data collected. This polemic, as is well known, resulted in a substantially negative verdict⁴ which can be summarized as follows: (a) the medieval MSS contain negligible differences with respect to the TR; (b) variants of any significance are rare and become submerged in the mass of allographies, copying errors, and secondary scribal innovations; (c) the contribution that medieval MSS can bring to the history and textual criticism of the HB is minimal: the uniformity of the medieval biblical text is better explained if we presume a single recension or a common archetype⁵ as the source of the current TR, and

⁴ For a historical account, see: D. BARTHÉLEMY, *Critique Textuelle de l'Ancien Testament. Josué-Esther*, I, Éditions Universitaires/Vandenhoeck & Ruprecht, Fribourg/Göttingen 1982, pp. 28-40; W. MCKANE, *Benjamin Kennicott: An Eighteenth-Century Researcher*, «The Journal of Theological Studies» 28 (1977), pp. 445-64; B. CHIESA, *Filologia storica della Bibbia Ebraica*, II, Paideia, Brescia 2000,

pp. 399-428; S. MANDELBROTE, *The Old Testament and Its Ancient Versions in Manuscript and Print in the West, from c. 1480 to c. 1780*, in E. CAMERON (ed.), *The New Cambridge History of the Bible. From 1450 to 1750*, Cambridge University, Cambridge 2016, pp. 82-109.

⁵ On the theories of a single recension or archetype, see M.H. GOSHEN-GOTTSTEIN, *Hebrew Biblical*

if we regard the variants of the MSS simply as the inevitable blemishes resulting from copying it.⁶

The thesis that the medieval MSS had a common origin negatively influenced the study of the medieval HB tradition according to normal philological criteria. What it has in fact led to is the automatic identification of all medieval MSS as *descripti* of the great Oriental Tiberian exemplars.⁷ As a result, not only has a stemmatic analysis of the data already available been sidelined, but even the idea of a new systematic collation of the MSS, to the point where, as Chiesa so rightly asserts: “after two hundred years we are

still stuck exactly at the first timid point, namely, the collations of Kennicott and De Rossi.”⁸

This failure to classify medieval MSS into families has in turn had consequences on ecdotic praxis. Thus, in the critical editions of the HB, the MSS of the two collations are either cited individually⁹ or simply counted – without providing any identifying *sigla*, and without distinguishing between MSS and printed editions¹⁰ – or indeed almost entirely excluded.¹¹

In the course of the last century, several attempts have been made to study the medieval HB tradition according to stemmatological crite-

Manuscripts. Their History and Their Place in the HUBP Edition, «Biblica» 48 (1967), pp. 253-64, and CHIESA, *Filologia*, pp. 420-6.

⁶ These are in effect the judgments expressed by scholars such as Eichhorn, Bauer, and Rosenmüller, for whom we refer to the work of Chiesa cited in note 4. In more recent times, this line of thought is especially typical of the Israeli school, see: GOSHEN-GOTTSTEIN, *Hebrew Biblical Manuscripts*; ID., *Biblical Manuscripts in the United States*, «Textus» 2 (1962), pp. 28-59; S. TALMON, *The Old Testament Text*, in P.R. ACKROYD - C.F. EVANS (eds.), *The Cambridge History of the Bible. From the Beginnings to Jerome*, Cambridge University, Cambridge 1970, pp. 159-99. Similar evaluations are expressed by the editorial board of the *Hebrew Old Testament Text Project* (HOTTP), see: D. BARTHÉLEMY, *Critique Textuelle de l'Ancien Testament. Ézéchiel, Daniel et les 12 Prophètes*, III, Éditions Universitaires/Vandenhoeck & Ruprecht, Fribourg/Göttingen 1992, pp. xix-xlix; J.A. SANDERS, *Text and Canon: Concepts and Method*, «Journal of Biblical Literature» 98 (1979), pp. 5-29; ID., *The Hebrew University Bible and Biblia Hebraica Quinta*, «Journal of Biblical Literature» 118 (1999), pp. 518-26.

⁷ This idea is polemically expressed by Goshen-Gottstein as follows: “I would consider a very provocative formulation, i.e., that in our case the *textus receptus* functions practically like the ‘archetype’ and that we might disregard the MSS by way of *eliminatio codicum descriptorum*. This sounds, of course, like a *reductio ad absurdum* of the Lachmannian method,” GOSHEN-GOTTSTEIN, *Hebrew Biblical Manuscripts*, p. 77, note 3.

⁸ CHIESA, *Filologia*, p. 414. The main argument against the actual usefulness of such an enterprise is date. The medieval MSS, especially if after the 13th century, would in fact be too late to guarantee the survival of real variants. The rule *recentiores non*

deteriores would lose almost all its force, in light of “the particular facts of the development of Hebrew Tiberian biblical MSS”, see GOSHEN-GOTTSTEIN, *Biblical Manuscripts*, pp. 31-2. On the application of this famous philological rule of thumb in HB studies, see C. MARTONE, *Recentiores non deteriores. A Neglected Philological Rule in the Light of the Qumran Evidence*, in F. GARCÍA MARTÍNEZ et al. (eds.), *4QMMT to Resurrection*, Brill, Leiden 2006, pp. 205-15.

⁹ This is the approach of the Italian school: both in studies of textual history and in critical editions, the use of the classical collations is systematic. For the former, see: B. CHIESA, *L'Antico Testamento ebraico secondo la tradizione palestinese*, Bottega d'Erasmus, Torino 1978; A. CATASTINI, *Isaia ed Ezechia. Studio di storia della tradizione di II Re 18-20 // Is. 36-39*, Università degli Studi ‘La Sapienza’, Roma 1989; ID., *L'itinerario di Giuseppe. Studio sulla tradizione di Genesi 37-50*, 1995. For the latter, see: P.G. BORBONE, *Il libro del profeta Osea. Edizione critica del testo ebraico*, Zamorani, Torino 1990; G. GARBINI, *Cantico dei Cantici. Testo, traduzione, note e commento*, Paideia, Brescia 1992; A. CATASTINI, *Storia di Giuseppe (Genesi 35-70)*, Marsilio, Venezia 1994.

¹⁰ As in the *Biblia Hebraica* series up to the fourth edition, the *Biblia Hebraica Stuttgartensia*, see also next note.

¹¹ This is the direction followed by more recent critical editions, such as the *Biblia Hebraica Quinta* (BHQ), the *American Hebrew Bible: A Critical Edition* (HBCE) and the *Israeli Hebrew University Bible Project* (HUBP). In the BHQ and in the HUBP the editors eschew, following Goshen-Gottstein, the use of traditional collations, limiting themselves to a few MSS judged relevant above all on the basis of external criteria such as antiquity, see A. SCHENKER et al. (eds.), *Biblia Hebraica Quinta. General Introduction and Megilloth*, Deutsche Bibelgesellschaft, Stuttgart 2004, p. xiv, and M.H. GOSHEN-GOTTSTEIN,

ria.¹² In most cases, however, the fundamental objective of such attempts has not so much been to arrive at a complete stemmatic classification, but rather to trace in the medieval documentation pre-Masoretic variants shared with the ancient Versions and therefore detect groups of MSS that are genealogically interrelated with these.¹³ Division into families is a mere function of the search for ancient variants and is consequently marginal – and with mostly negative results¹⁴ – or absent altogether.

Two contributions, more recent and articulated in terms of method, are exceptions: that of Sacchi (Genesis) and of Borbone (Hosea).¹⁵ Exploiting the computer, these scholars propose a path towards a computational treatment of variants and grouping of witnesses based on the criterion of distance.¹⁶

The Book of Isaiah. Sample Edition with Introduction, Magnes, Jerusalem 1965. In the HBCE the editors cite the MSS only desultorily and solely in the case of variants considered significant, see: R. HENDEL, *The Oxford Hebrew Bible: Prologue to a New Critical Edition*, «Vetus Testamentum» 58 (2008), pp. 324-51, and M.V. FOX, *Proverbs. An Eclectic Edition with Introduction and Textual Commentary*, Society of Biblical Literature, Atlanta 2015, pp. 17-8.

¹² A review of these studies can be found in BARTHÉLEMY, *Critique textuelle*, pp. xix-xxvii. These are, in chronological order: J. HEMPEL, *Chronik*, «Zeitschrift für die alttestamentliche Wissenschaft» 48 (1930), pp. 187-206; Id., *Innemasoretische Bestätigungen des Samaritanus*, «Zeitschrift für die alttestamentliche Wissenschaft» 52 (1934), pp. 254-74; J.W. WEVERS, *A Study in the Hebrew Variants in the Books of Kings*, «Zeitschrift für die alttestamentliche Wissenschaft» 61 (1948), pp. 43-76; M.H. GOSHEN-GOTTSTEIN, *Die Jesaiah-Rolle und das Problem der hebräischen Bibelhandschriften*, «Biblica» 35 (1954), pp. 429-42; H. GESE, *Die hebräischen Bibelhandschriften zum Dodekapropheten nach der Variantensammlung des Kennicott*, «Zeitschrift für die alttestamentliche Wissenschaft» 69 (1957), pp. 55-69. To these must be added the studies cited at note 15, as well as the research of two pupils of Goshen-Gottstein, M. Cohen and J. Penkower, focused on the history of the medieval manuscript tradition and on the relationship between this and the first printed editions, in particular the Rabbinic Bibles: M. COHEN, *מגבשי כתיב במצחפי מסורה*, עתיקים ומשמעם לתולדות נוסח המקרא המקובל, Hebrew University, Jerusalem 1973 Ph.D. diss. [not pub-

lished]; Id., *האידיאה בדבר קדושת הנוסח לאותיותיו וביקורת*, הטכסט, in U. SIMON (ed.), *המקרא ואנחנו*, I, דביר, Tel Aviv 1979, pp. 42-69; Id., *קווי יסוד לדמותו העיצורית של*, הטכסט בכתבי יד מקראיים ממי הביניים, in U. SIMON - M.H. GOSHEN-GOTTSTEIN (eds.), *עיוני מקרא ופרשנות*, I, Bar Ilan University, Ramat Gan 1980, pp. 123-82; Id., *מהו "נוסח המסורה"*, ומה היקף אחיזתו בתולדות המסירה של ימיה"ב, in U. SIMON (ed.), *עיוני מקרא ופרשנות*, II, 1986, pp. 229-56; Id., *יעקב בן חיים וצמיחת מהדורת המקראות*, הגדולות, Hebrew University, Jerusalem 1982, Ph.D. diss. [not published].

The effectiveness of such a method is also proven by Borbone, who studied the book of

lished]; Id., *האידיאה בדבר קדושת הנוסח לאותיותיו וביקורת*, הטכסט, in U. SIMON (ed.), *המקרא ואנחנו*, I, דביר, Tel Aviv 1979, pp. 42-69; Id., *קווי יסוד לדמותו העיצורית של*, הטכסט בכתבי יד מקראיים ממי הביניים, in U. SIMON - M.H. GOSHEN-GOTTSTEIN (eds.), *עיוני מקרא ופרשנות*, I, Bar Ilan University, Ramat Gan 1980, pp. 123-82; Id., *מהו "נוסח המסורה"*, ומה היקף אחיזתו בתולדות המסירה של ימיה"ב, in U. SIMON (ed.), *עיוני מקרא ופרשנות*, II, 1986, pp. 229-56; Id., *יעקב בן חיים וצמיחת מהדורת המקראות*, הגדולות, Hebrew University, Jerusalem 1982, Ph.D. diss. [not published].

¹³ On this question, see the penultimate paragraph of § 6, and especially notes 67 and 69.

¹⁴ See WEVERS, *A Study*, p. 75: "it is almost, if not wholly, impossible to classify the Hebrew mss. by groups or families. This is what one might expect since the copyists of these mss. were undoubtedly influenced by M [Masoretic text]. In consequence, the scribes brought their mss. more and more into the Masoretic tradition [...]."

¹⁵ P. SACCHI, *Analisi quantitativa della tradizione medievale del testo ebraico della Bibbia secondo le collazioni del De Rossi*, «Oriens Antiquus» 12 (1973), pp. 1-13; BORBONE, *Osea*, pp. 183-227.

¹⁶ A complete description of the method can be found in SACCHI, *Analisi quantitativa*, pp. 6-8. As far as we know, Sacchi was the first to employ a computer-assisted method for the stemmatic analysis of the medieval tradition of the HB.

¹⁷ This is group β, with a remarkable 225 MSS, of which 36 exclusive to the group, see SACCHI, *Analisi quantitativa*, pp. 9ff.

¹⁸ SACCHI, *Analisi quantitativa*, pp. 8ff.

Hosea by using the data provided by KN. In this case, however, the groupings produced do not stand up to qualitative analysis: while the mathematical method does allow for division of the MSS into groups, the high number of graphic variants on which they are based makes such a division useless for purposes of *eliminatio descriptorum* and, more generally, for textual criticism.¹⁹

Barthélemy arrived at a similar conclusion after the termination of the third volume of his *Critique Textuelle*: for him, it is impossible to identify families of codices, at least not by resorting solely to the data contained in traditional collations.²⁰ The medieval tradition, he asserts, can only be described in terms of codices copied with a more or less variable accuracy, and subsequently subjected to recension.²¹ According to this view, one might conclude, variants cannot be considered indicative of genetic kinship, but as mere scribal idiosyncrasies.²²

As things stand now, the thesis seconded by Barthélemy predominates in the literature, while Sacchi's stemmatic attempt has – as far as we know – generated no follow-up.

It is precisely from the antithesis of these two positions regarding the feasibility of a stemmatic reconstruction that our research takes its jumping-off point. Using Q as an example, and generalizing, with due caution, from the results of our inquiry into this one biblical book, we will

attempt to verify whether a stemmatic analysis of the medieval tradition is inevitably doomed to failure, or, on the other hand, if it might indeed be possible to identify textual groups or families, and if so, to what extent.

Let us move on now to present the data (§ 3) and the method (§ 4) and to discuss how we have adapted them for the purpose of stemmatic analysis.

3 The Evidence

3.1 Witnesses

As indicated in the Introduction, we have considered the witnesses of KN, but with important limitations, as well as additions. To avoid fragmentary results, we have excluded those MSS which in KN appear as only partially collated, as well as those MSS which, although collated in full, do not contain the entire text of Q (nos. 58 and 129). Very late MSS (145, 147, 246, 248, 249, and 251, all 18th-century) and a *perush* (57) were also excluded.

Of the 16 printed editions collated by KN, we have retained only the reference text, namely, the edition of E. van der Hooght (Amsterdam 1705), plus six others judged significant for the history of the text (259, 260, 264, 270, 271, 275).

¹⁹ BORBONE, *Osea*, p. 192.

²⁰ BARTHÉLEMY, *Critique textuelle*, p. xxvii.

²¹ BARTHÉLEMY, *Critique textuelle*, p. xxxii.

²² Beyond generic affirmations regarding a generalized contamination of the textual tradition, Barthélemy does not adduce positive evidence to support the thesis of the impossibility of dividing the Hebrew codices into families. The only argument cited is the results of the investigation on Hosea made by Borbone: “[the] impossibility of grouping into textual families the manuscripts that diverge from the TR seems to constitute yet another important finding implicit in Borbone’s quantitative analysis” (p. xxvii). In fact, Borbone explicitly states that the method leads to the identification of groups (p. 192); that these are based almost exclusively on graphic variants is a consequence of using KN as a whole, and does not constitute direct evidence; in addition, the survey is conducted on only 11 of Hosea’s 69 MSS (24 if the fictitious MSS generated by scribal interventions are included, see § 3.2) and is therefore

deliberately partial. When he turns his attention to Sacchi’s work, by contrast, Barthélemy sidesteps altogether the division into groups proposed there, claiming that DR is not reliable (see also note 67). Even Goshen-Gottstein, who among HB scholars is perhaps the most reluctant to attribute any importance to the study of medieval codices, leaves the door open on this front, when he writes: “[n]o stemmatic typology of Hebrew medieval manuscripts had ever been attempted after the Kennicott debacle. The huge number of medieval Hebrew biblical manuscripts are usually referred to by the numbers given by eighteenth-century collators, since stemmatic attempts are seemingly doomed. [...]. *This does not mean that certain individual manuscripts could not be judged as stemmatically connected to one another*” (emphasis added), M.H. GOSHEN-GOTTSTEIN, *The Textual Criticism of the Old Testament: Rise, Decline, Rebirth*, «Journal of Biblical Literature» 102 (1983), p. 394, note 103.

We have added nine MSS collated *ex novo* based on criteria such as antiquity, availability, and completeness of the text. Important but fragmentary MSS were not taken into account.²³

Among the nine, there are 3 MSS that KN collated partially (326, 590, 602) and six which are altogether absent in his catalogue. Here is the list:²⁴

<i>Signum</i>	Century	Script	Identifier
1-ASS82	1189	Ashkenazi	Sassoon, David Solomon, London, Ms. 282; Museum of the Bible, Washington, Ms. 858
2-AAdd9403	XIII	Ashkenazi	Add. 9403, The British Library, London
0-OEVRIIB55	XI	Oriental	Ms. EVR II B 55, The National Library of Russia, St. Petersburg
0-OEVRIIB94	XI	Oriental	Ms. EVR II B 94, The National Library of Russia
0-OEVRIIB34	XI	Oriental	Ms. EVR II B 34, The National Library of Russia
0-OL	XI	Oriental	Firkovich B 19 A (<i>Codex Leningradensis</i>)

3.2 Variants

With regard to variants, too, certain limitations have been necessary. Since we set as our primary objective the individuation of *genealogical* relationships among MSS, the problem of distinguishing between mono- and polygenetic variants arose. To this end, we devised a classification scheme that allowed us to describe variants on the basis of typological criteria and to

identify, on the basis of these criteria, only those with greater kinship-revealing power.

The main distinction we have applied is between substantial and accidental variants.²⁵ Among the former, we include the phenomena of addition, deletion, substitution, and transposition, while among the latter are variants related to spelling in general (such as those of *matres lectionis* and certain instances of graphic variants²⁶) and to word division.²⁷ The former are

²³ Fragmentary witnesses are more difficult to deal with, both at the level of variant encoding and at the computational level. For this reason we have decided to exclude them.

²⁴ In order to facilitate the reading of the graph in Fig. 4, we have created a special notation for the *sigla* of the Kennicottian witnesses. For MSS, the century is indicated by a number always placed at the beginning ('0' for the 11th century, '1' for 12th century etc.); the script is indicated by a capital letter ('A' for Ashkenazi, 'S' for Sephardi, 'O' for Oriental, 'I' for Italian and 'IA' for Italian-Ashkenazi); the catalogue number follows ('K1', 'K2' etc.). Printed editions, on the other hand, are indicated by the letter 'E' plus the catalogue number. The edition of van der Hooght is indicated by 'H'. The information on script and date is taken, when present, from the online catalogue of the National Library of Israel (<https://web.nli.org.il/sites/nlis/en/manuscript>), otherwise from DR and, if missing there as well, from KN. If neither the script nor the date is known, a MS is identified by the symbol 'U' (unclassified) and by a question mark respectively.

²⁵ On the typological distinction between substantial and accidental variants, see G. CONTINI, *Breviario di ecdotica*, Einaudi, Torino 1990, pp.

38ff. According to Contini, the first to have theorized this difference in terms of *critique des formes* and *critique des leçons* was Gaston Paris, in his introduction to the edition of the Life of Saint Alexis (1872). A more recent formulation of this theory is found in W. W. GREG, *The Rationale of Copy-Text*, «Studies in Bibliography» 3 (1950), pp. 19-36. In Greg's formulation, the distinction between substantial and accidental is adopted by the editorial board of the HBCE as a rationale for the spelling, vocalization, and punctuation of the critical text, see: R. HENDEL, *Prologue*; ID., *The Oxford Hebrew Bible: Its Aims and a Response to Criticisms*, «Hebrew Bible and Ancient Israel» 2 (2013), pp. 63-99; FOX, *Proverbs*, pp. 19ff.

²⁶ By graphic variants, we mean: (1) allographies that do not involve a change of meaning (for example the different forms of the relative pronoun, the demonstrative, and some personal pronouns; nouns such as לב/לבו, אר/אריה etc.); (2) allographies that can affect the meaning (minimal pairs such as שכלות/סכלות 'wisdom' or 'wisdom/foolishness', also attested in the ancient Versions); (3) errors (for example those relating to the interchange or the loss of the gutturals). On the treatment of these variants, see next note.

²⁷ Distinguishing between substantial and accidental variants is not always straightforward in

more likely to be significant from a genealogical standpoint and thus are those on which we base our analysis.²⁸ The latter, on the other hand, are more easily attributable to chance or to particular scribal habits and can constitute a problem for quantitative analysis due especially to their numerical preponderance.²⁹

We have also eliminated from consideration those readings occurring in a single codex (*lectiones singulares*), both substantial and accidental, as they are useless for defining relationships between MSS.³⁰

Hebrew, especially in the absence of disambiguating vocalization. See the criticisms leveled at the HBCE editorial policy referred to in the previous note: H.G.M. WILLIAMSON, *Do We Need a New Bible? Reflections on the Proposed Oxford Hebrew Bible*, «Biblica» 90 (2009), pp. 153-75, and E. TICHELAAR, *Editing the Hebrew Bible: An Overview of some Problems*, in J.S. KLOPPENBORG - H.N. JUDITH (eds.), *Editing the Bible: Assessing the Task Past and Present*, Society of Biblical Literature, Atlanta 2012, pp. 41-65. In some cases it may be difficult to choose between a variant of *mater lectionis* or graphic variant and a morphological or semantic substitution. For example, consider the pairs ירא/ירא, ידך/ידך etc., interpretable both as alternating forms of *scriptio plena/defectiva* and as distinct lexemes/morphemes (respectively, ‘fear/see’, ‘your hand/your hands’). The same goes for graphic variants such as שכלות/סכלות mentioned in the previous note. Faced with such ambiguities, we have followed two different strategies: as to the variants of *matres lectionis*, given their high number, we have always classified them as accidental and therefore have excluded them; as to the graphic variants, being rarer, we have proceeded case by case on the basis of qualitative considerations. Indeed, some graphic realizations can be judged genealogically relevant, for example האסורים/הסורים in Q 4,14. For most of the graphic variants, however, the problem of a correct typological classification is of lesser importance for two reasons: (1) they are a tiny minority among the accidental variants (4%, see note 29); (2) most of them are *lectiones singulares*.

²⁸ It is clear that the substantial variants remaining after regularization cannot be considered monogenetic in the Lachmannian sense (§ 4): many concern particles such as the article, conjunctions, and prepositions, that is, minimal parts of speech frequently subject to accidental variation. In dealing with such cases we found ourselves faced with three possibilities: (1) to exclude everything that cannot be

The treatment of first- and second-hand variants merits separate discussion. As is well known, biblical MSS usually present a certain number of scribal interventions. Some of these were intended to eliminate original variants to make the text conform to the vulgate; others actually eliminated the readings of the vulgate by generating new variants. The first are those that KN indicates with the Latin term *primo*, while the second are those introduced by *nunc*.³¹ For each MS it is therefore possible to hypothetically distinguish two types of text: one prior to the

considered strictly monogenetic; (2) to elaborate a system of weighting factors, assigning to each variation phenomenon a numerical value corresponding to the kinship-revealing power (for example, minimal for the variants concerning particles and maximum for the root substitutions); (3) to work with all substantial variants without distinction, trusting that the cases of random variation will be exceeded by the number of genetic variants. Among the three, we chose the last route. The first would, *inter alia*, risk reducing the database excessively, making it unsuitable for statistical investigation (see on this also note 49). The second would involve the introduction upstream of the analysis of significant qualitative assumptions, which would be added to those already used to distinguish between substantial and accidental variants. In fact, we believe that this distinction is already sufficient for the purpose we set ourselves here – that is, to suggest a way to manage the plethora of medieval variants – and that our method represents a good compromise between the qualitative principles of philological method and the minimum requirements of quantitative analysis.

²⁹ 63% of total variants are accidental. Of these, approx. 96% are represented by variants of *scriptio plena/defectiva*.

³⁰ See for example P. MAAS, *Textkritik*, Teubner, Leipzig 1952², p. 7, and W.W. GREG, *The Calculus of Variants. An Essay on Textual Criticism*, Oxford 1927, p. 19. In biblical studies this point was remarked by B. CHIESA, *Textual History and Textual Criticism of the Hebrew Old Testament*, in J.C. TREBOLLE BARRERA (ed.), *The Madrid Qumran Congress I*, Brill, Leiden 1992, p. 267. The algorithm that we use, described in § 4, excludes by default the *lectiones singulares* as non-informative, see P. ROELLI (ed.), *Handbook of Stemmatics. History, Methodology, Digital Approaches*, De Gruyter, Berlin/Boston 2020, pp. 311-2. In Q, about 40% of the substantial variants are made up of *lectiones singulares*.

³¹ See DE ROSSI, *Variae Lectiones*, I, p. xlvi:

correction and one subsequent. In concrete terms, in the first case all the *primo* variants will be examined for each MS, with the *nunc* variants excluded, and *vice versa* in the second.

The literature sometimes has it that the different hands of a copyist should be treated as separate MSS.³² Thus, if a MS *A* contains corrections, we would have a MS *A*¹ containing the text without corrections (i.e., the original, first-hand text) and a MS *A*² containing the text resulting from these corrections (i.e., the corrected, second-hand text). This method has undoubted advantages because it enables one to identify, for example, the antigraph from which the correctors drew the corrections. More generally, it makes it possible to distinguish the variants generated by different actors at different moments in the history of the text, thus avoiding confusion in the reconstruction of the transmission process.

In our situation, however, this method is unsuitable for two reasons. One has to do with the reality of the transmission of the medieval biblical text: many MSS have in fact too few variants and very few corrections. Dividing such MSS into two separate entities would not be useful, either for the purposes for which this method was conceived or in general.³³ As Sacchi suggests, the most significant occurrences can be better studied in the context of qualitative analysis.³⁴

The second, more important reason is related to the quality of the data in our possession. KN, as is known, does not consistently indicate from which hand of the copyist the variant originated.³⁵ From the re-collation which we personally conducted of fifty-nine MSS, it emerged that only 11% of such variants are correctly annotated.³⁶

However, even if the lack of reliable data relating to scribal interventions does not allow us to apply the foregoing method, it does not constitute an impediment to analysis. Indeed, only the corrections that generate variants – those annotated with *nunc* – represent a problem from a stemmatological point of view. Such corrections may in fact derive from the consultation of other sources such as MSS or Masoretic lists – that is, they can be the product of contamination.³⁷ Fortunately, the possibility of these variants distorting the results is limited for two reasons: in general, they are low frequency variants,³⁸ and most of the time KN omits them entirely.³⁹

The problem does not arise, by contrast, in the case of first-hand variants (*primo*). As a rule, KN reports them, but without specifying that they are subsequently-corrected variants – that is, without adding the annotation *primo*. This omission is of little importance to us: as we have said, these variants are *always* corrected in conformance with the vulgate. The fact that

“[q]ui ergo codices praefixum prius vocabulum habent [i.e. *primo*], ita olim legisse intelliguntur, sed nunc emendati et ad masoram reformati, contra qui posterius [i.e. *nunc*], antea hodiernam lectionem habuisse a posteriori manu in variantem mutatam.”

³² For example in J. FROGER, *La collation des manuscrits à la machine électronique*, «Bulletin d'information de l'Institut de Recherche et d'Histoire des Textes» 13 (1964), pp. 135-71, cited in SACCHI, *Analisi quantitativa*, p. 7. The method to which Sacchi refers is implemented in BORBONE, *Osea*, pp. 185ff.

³³ To this consideration can be added a practical one: Since the majority of MSS have at least one first- or second-hand variant, there is a risk not only of doubling the number of witnesses, thus making it difficult to interpret the results, but also of obtaining specious associations: a pair of MSS generated after splitting could result in fact as associated, being essentially the same text except for (mostly a very few) scribal interventions.

³⁴ SACCHI, *Analisi quantitativa*, p. 7.

³⁵ This fact was already clear to DR, who on this point expressly promised greater precision, see DE ROSSI, *Variae lectiones*, I, pp. xlvi-xlvii.

³⁶ The estimates that we provide in the rest of this section are calculated on the small sample of MSS reviewed and refer only to the substantial variants. The measure used is that known as accuracy, calculated as the ratio between the cases recorded in KN and the total of the variants present in the MSS according to our findings.

³⁷ As Cohen has shown, see COHEN, קווי יסוד. See also BARTHÉLEMY, *Critique textuelle*, p. xxxii.

³⁸ The second-hand variants represent only 3% of the total, while the first-hand ones are much more numerous (41%). This can be easily verified by glancing at the column of the *nunc* variants in Appendix A.

³⁹ Only in very rare cases does KN report them without specifying that they are subsequent interventions.

they are first-hand variants is irrelevant for purposes of stemmatic analysis.

In light of all of these considerations, we have examined first-hand variants but eliminated those of the second hand, in order to limit the impact of contamination.⁴⁰ For the same reason we have also excluded marginal variants.⁴¹ Dubious variants, both first- and second-hand, as well as readings over erasures,⁴² are equally excluded.⁴³

Overall, the data we have collected can be considered reliable, even for those MSS that we have not been able to double-check. In fact, if we exclude those scribal interventions for which, as mentioned, KN is not reliable, the general accuracy of the collation can be estimated at approximately 86%.⁴⁴ This means that the cases in which KN fails to mention substantial variants are, after all is said and done, relatively few. Even rarer are the cases in which a variant is erroneous or wrongly attributed to a given MS. Finally, most of the unchecked MSS are among those that have a

high number of variants in KN and that therefore the collators have scrutinized thoroughly: even if it cannot be excluded entirely that some variant or other has escaped or been encoded erroneously, this is unlikely to significantly affect the results we are about to present.

The total number of witnesses examined is 116. The variant locations, the places in the text that present variants, number 371. To allow processing by the machine, both witnesses and variant locations have been arranged in a data structure called a matrix (Fig. 1).⁴⁵ In the matrix, witnesses are represented as rows of a table, while variant locations are represented in the columns. Variants are encoded as follows: if in a given variant location a witness does not have any variant – that is, if in that particular point it is identical to the reference text – we assigned ‘0’; otherwise, if a variant is attested, we assigned a number according to the order in which it appears in the critical apparatus: ‘1’ if it is the first variant, ‘2’ the second, and so on.

⁴⁰ Contamination is a natural part of a textual tradition and not taking it into account inevitably leads to a loss of information. In our case, however, the choice to exclude potentially contaminating readings – and to act, in fact, as if the contamination does not exist – is an obligatory simplification dictated by the method of analysis chosen: as is well known, the genealogical method – and, more precisely, the tree model underlying it – is not designed to directly manage contamination. For a recent treatment on this problem, see ROELLI (ed.), *Handbook of Stemmatics*, pp. 254-72; for the specific limits of the phylogenetic method we use (§ 4), see C. HOWE *et al.*, *Responding to Criticism of Phylogenetic Methods in Stemmatics*, «Studies in English Literature» 52 (2012), pp. 56-8.

⁴¹ These too may be the result of contamination, as KN already noted: “[i]n margine codicum Hebraeorum saepenumero notantur Variarum Lectiones; ex aliis codicibus vel diserte, vel probabiliter, petita,” KENNICOTT, *Vetus Testamentum*, I, p. iii.

⁴² These are divided into two categories: those annotated with *sup. ras.* and those characterized by a symbol that imitates the erasure, roughly ‘#’ (see KENNICOTT, *Vetus Testamentum*, I, p. ii). Both are unusable for us because the identification of the variant is either doubtful or impossible. In the first case, the annotation *sup. ras.* only indicates the presence of an erasure upon which the vulgate reading was later superimposed (in only one instance out

of about forty did we find an example of a genuine variant, in Q 3,5); in the second case, the presence of the erasure – when not simply due to a physical blemish on the scribal support (e.g. Q 2,8) – obscures completely the original variant, which can therefore only be conjectured (as DR does in some cases, see MSS K82 and K158 in Q 8,3).

⁴³ Both are of an uncertain nature, and therefore are to be eliminated. The former, as the Latin annotation that identifies them (*videtur, forte*) indicates, are dubious and could also be reading errors on the part of the collators. The latter are often not even actual variants (see previous note). The case of readings over erasures – not to mention that of *voces imperfectae* (KENNICOTT, *ibid.*), another category of ‘non-variants’ also excluded – is indicative of a rather ‘impressionistic’ way of collating, which confirms the judgment of Chiesa about the lack, in KN, of *the very concept of variant*, see CHIESA, *Filologia*, p. 411.

⁴⁴ The estimate drops sharply if the first- and second-hand variants are taken into account: 46%.

⁴⁵ The matrix is calculated from an .xml file encoded according to the standards of the *Text Encoding Initiative* (<https://tei-c.org/>). Details on the process of digitizing variants and xml encoding can be found in L. BAMBACI, *Critical Apparatus as Domain Specific Languages. A Rule-Based Parser for Encoding an Eighteenth-Century Collation of Hebrew Manuscripts*, «International Journal of Information Science and Technology» 5 (2021), pp. 22-33.

In the following section we will explain the use of the matrix for the calculation of genealogical relationships. For a complete list of witness-

es, please refer to Appendix A at the end of the article.

Taxon	1	2	3	4
H	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
00L	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
3IAK1	000000000000000001000100000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
3SK2	000010000000100000010000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
4SK3	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
2AK4	0000010100000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
4SK14	0010000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
3AK17	000001000001000000010010000000000010100	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
3AK18	0100001000010000100110100000000101010000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
4SK19	000000000001000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
2AK30	000000100010000001000010000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
00K31	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000	0000000000000000000000000000000000
...				

Figure 1: Fragment of data matrix.

4 Method

Before commencing our quantitative analysis, we attempted to trace any monogenetic errors in KN’s apparatus, to see if it was possible to apply Lachmann’s genealogical method.⁴⁶ The results of this investigation were disappointing: none of the consonantal variants of q can be considered monogenetic in the strict sense, that is, such that they cannot be explained by polygen-

esis (criterion of irreproducibility) or corrected by a scribe by conjecture (criterion of irreversibility).⁴⁷ The former, such as variants of *matres lectionis* and other similar *minutiae*, are the vast majority and have little or no conjunctive power; the latter, such as obvious errors, are mostly *lectiones singulares* and therefore equally useless for establishing witness relationships.

In the face of a situation like this,⁴⁸ the application of quantitative criteria to textual anal-

⁴⁶ This is the best-known genealogical method in philology. In the realm of textual criticism of the HB, however, it remains marginal, with the exception of certain studies and ‘eclectic’ editions (see for example the bibliography at notes 9 and 11). The first modern study based on Lachmannian methodology is P. SACCHI, *Il rotolo A di Isaia. Problemi di storia del testo*, «Atti dell’Accademia Toscana di Scienze e Lettere ‘La Colombaria’» 30 (1965), pp. 31-111. The credit for having called the attention of scholars to the importance of the method, however, goes to the famous article by Chiesa cited at note 30, directed precisely against the statistics-based quantitative methods used in Qumran studies: CHIESA, *Textual History*; see also Id., *Il Testamento Biblico. Condizioni antiche e moderni*, in P. SACCHI (ed.), *Il giudaismo palestinese: dal I secolo a.C. al I secolo d.C.*, Atti dell’VIII congresso internazionale dell’AISG, San Miniato, 5-6-7 novembre 1990, Fattoadarte, Bologna 1993, pp. 5-18. For a recent example of application of Lachmannian criteria, see C. MARTONE,

Separative and Conjunctive Errors: A Note on a New Reading of 4Q99 23, «Henoch» 42 (2020), pp. 391-5. On the *status quaestionis*, see: A. CATASTINI, *Da Qumran al Testamento Masoretico dell’Antico Testamento: Spunti metodologici per la valutazione delle varianti*, «Revue de Qumrân» 15 (1991), pp. 303-13; CHIESA, *Filologia*, pp. 432ff.; C. MARTONE, *Il Giudaismo antico. 538 a.e.v.-70 e.v.*, Carocci, Roma 2008, pp. 106ff.

⁴⁷ See C. MACÉ et al., *Comparing Stemmatalogical and Phylogenetic Methods to Understand the Transmission History of the Florilegium Coislinianum*, in A. BUCOSSI - E. KIHLMAN (eds.), *Ars Edendi Lecture Series*, II, Stockholm University, Stockholm 2012, p. 115. On the minimal requirements of Lachmann’s method, see also M. WEITZMAN, *The Analysis of Open Traditions*, «Studies in Bibliography» 38 (1985), p. 92.

⁴⁸ That a qualitative examination of the apparatus of classical collations is doomed to failure, whatever the biblical book, seems also proved by the at-

ysis seems the only practicable one and finds its ultimate justification in an argument based on ordinary probability: even if an agreement in polygenetic reading has no value in itself because it may be due to mere chance, the likelihood that more such agreements are random is lower the greater their number.⁴⁹

The second and principal phase of our analysis was therefore quantitative in nature. To implement it, we resorted to the techniques of phylogenetic systematics, namely, that branch of biology which specializes in the study of the evolution of living organisms.⁵⁰

Over the years, phylogenetic techniques in the field of stemmatology have acquired a certain history, given the vast corpus of phylogenetic studies of all sorts of manuscript traditions.⁵¹

tempts of Borbone and Catastini, see BORBONE, *Osea*, p. 192, and CATASTINI, *L'itinerario di Giuseppe*, p. 334. See also the words of Chiesa in this context, cited in the next note.

⁴⁹ In Varvaro's words: "with the same weak conjunctive force, would it not be legitimate to consider a long series of errors more valid than a small one? In fact, the weakly conjunctive error is intrinsically polygenetic, so that in theory both the short series and the more ample one could be random, but it is evident that this is less likely the longer the series. It is true that in textual criticism we should not take into account elements that could be fortuitous, but if there are no others we must put to good use those we have," A. VARVARO, *Critica dei testi classica e romanza. Problemi comuni ed esperienze diverse*, «Rendiconti dell'Accademia di Archeologia, Lettere e Belle Arti di Napoli» 45 (1970), p. 95. These considerations are certainly valid in the case of the medieval tradition of the HB. Hence the importance of including variants traditionally considered of little significance, such as those relating to particles (see note 28). See on this point the words of Chiesa regarding the variants of addition or omission of the copulative conjunction: "if these variants are considered insignificant, it would be opportune to revise the concept of variant; but it is equally appropriate to remember that, until proven otherwise, the transmitted text should be the same, so that the macroscopic variants cannot represent the norm, in a text which, moreover, did not have an exclusively literary life, but was inevitably subject to a certain control, due to the religious use for which it was intended," CHIESA, *L'Antico Testamento*, pp. 315-6.

⁵⁰ In the remainder of this section, only the ba-

This interaction between genealogical textual criticism and evolutionary biology was made possible, at least from the early 1990s, by virtue of both disciplines' sharing a genealogical-evolutionary model.⁵² Unlike the classic Lachmann approach, however, phylogenetic methods presuppose neither the selection of monogenetic errors, which is inevitably subjective and, as we have seen, often not even viable, nor the usual distinction between original reading and error, which is central to the Lachmannian paradigm. Instead, they exploit all of the variants – or at least a subset of significant ones, as in our case here – whether original or not.

As a rule, phylogenetic analysis of a textual tradition progresses in two stages: to begin with, variants are utilized to reconstruct the

principles of the method will be illustrated. The technicalities are deliberately reduced to the minimum possible. For a more detailed discussion, we refer to L. BAMBACI, *Digitizing Kennicott's Collation of the Hebrew Bible. Experiences of Encoding and of Computer-Assisted Stemmatic Analysis* [forthcoming]. For a general introduction to phylogenetic analysis, see R.D.M. PAGE - E.C. HOLMES, *Molecular Evolution. A Phylogenetic Approach*, Blackwell, Oxford 1998.

⁵¹ As regards applications in philology, the classical works are: A.R. LEE, *Numerical Taxonomy Revisited: John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum*, «Studia Patristica» 20 (1989), pp. 24-32; P. VAN REENEN - M. VAN MULKEN (eds.), *Studies in Stemmatology*, I, Benjamins, Amsterdam/Philadelphia 1996; B.J.P. SALEMANS, *Building Stemmas with the Computer in a Cladistic, Neo-Lachmannian, Way. The Case of Fourteen Text Versions of Lanceloet Van Denemerken*, Nijmegen University, 2000, Ph.D. diss.; P. VAN REENEN *et al.* (eds.), *Studies in Stemmatology*, II, Benjamins, Amsterdam/Philadelphia 2004. A general overview can be found in the most recent ROELLI (ed.), *Handbook of Stemmatology*, chapter V.

⁵² On this interaction, see: N.I. PLATNICK - H.D. CAMERON, *Cladistic Methods in Textual, Linguistic, and Phylogenetic Analysis*, «Systematic Biology» 26 (1977), pp. 380-5; R.J. O'HARA, *Trees of History in Systematics and Philology*, «Memorie della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale di Milano» 27 (1996), pp. 81-8; C. HOWE *et al.*, *Parallels Between Stemmatology and Phylogenetics*, in P. VAN REENEN *et al.* (eds.), *Studies in Stemmatology*, II, pp. 3-11. For an attempt at a

‘chain’ of relationships among witnesses; then, and only then, a temporal order is imposed on this chain by means of a selection of (a few) original variants. The graph obtained in the end, called a phylogram, constitutes for all intents and purposes a hypothesis as to the history of transmission, and is not dissimilar, in this respect, to a traditional *stemma codicum*. Maintaining a distinction between these two moments – construction of the chain and, thereafter, its orientation – is central to this method and implies a considerable advantage, at the operative level, over the traditional philological approach: In point of fact, while in Lachmann the absence of errors renders impossible the arrangement of a stemma and therefore blocks the road to any further inquiry, in phylogenetic analysis the use of variants instead of errors allows us at least to explore a textual tradition via the reconstruction of genealogical relationships, which is precisely the goal we have set for ourselves here.⁵³

In phylogenetic analysis, genealogical relationships are represented by a type of graph called a phylogenetic tree. In a phylogenetic tree, objects (*taxa*) are represented as terminal nodes (leaves), while the relationships among *taxa* are represented by the branches. These relationships are computed thanks to the alignment of the characters – in this case, the variant locations – in a matrix similar to the one presented above, and are expressed in terms of *genealogical descent*: related *taxa* are gathered together into the same group having as its progenitor a hypothetical *taxon* (intermediate node or internode), from which it is assumed that they have inherited their respective common traits, as we are about to discuss.⁵⁴

comparative history of the genealogical method in both biology and philology, see W. ROBINS, *Editing and Evolution*, «Literatur Compass» 4 (2007), pp. 89-120.

⁵³ The supporters of phylogenetics have staunchly insisted on this point, see for example SALEMANS, *Building Stemmas*, p. 5, and ROBINS, *Editing and Evolution*, p. 92. The need to establish from the outset the *direction* of textual change for the purpose of constructing a stemma represents one of the major problems of Lachmann’s method, see D.C. GREETHAM, *Textual Scholarship. An Introduction*, Garland, New York/London 1994, pp. 323 ff.

There exist several bioinformatic algorithms for creating phylogenetic trees. Among these we have selected the one that implements the principle of *maximum parsimony*. According to this principle, it is more economical to explain the relatedness of two *taxa* by hypothesizing a common descent of the shared characters, rather than an independent origin. In other words, it is more probable that two *taxa* are related because they descend from a common ancestor, while it is more difficult – less parsimonious – to hypothesize that they have developed similar characters by other means, such as by convergence (or parallelism) and reversion⁵⁵ – these latter phenomena roughly corresponding in philology to polygenesis and correction by conjecture. Such phenomena, collectively subsumed in biology under the general rubric *homoplasia*, require in fact the introduction of a greater number of changes in the reconstruction of a scheme of genetic relations.

To illustrate, suppose we have four MSS *ABCD* and a variant location with two possible states: ‘0’ for *AB* and ‘1’ for *CD*. The simplest, most parsimonious, hypothesis is that *AB* and *CD* inherited their respective variants from two common ancestors. If we assume that the reading of the archetype is ‘1’, we will have only a single change of state, namely, the one produced by the passage from the archetype to the hyparchetype of *AB*, as depicted by the phylogenetic tree in Fig. 2.

If, instead, we suppose that *AC* descend from a common hyparchetype, then we are compelled to imagine two independent changes of state: one for *A* and one for *B*, as shown in Fig. 3. This reconstruction requires an assumption of homoplasia and is therefore to be discarded as not parsimonious.

⁵⁴ This is an important difference with respect to the clustering algorithms used in computational stemmatology, such as that of Sacchi. These, in fact, group the witnesses on the basis of the *total* of variants through the calculation of various distance metrics; phylogenetic methods, by contrast, analyse the *evolution* of the individual variants, taking into consideration only the inherited ones, see R.J. O’HARA - P. ROBINSON, *Report on the Textual Criticism Challenge 1991*, «Bryn Mawr Classical Review» 3 (1992), p. 334, and LEE, *Numerical Taxonomy*, p. 26.

⁵⁵ See PAGE - HOLMES, *Molecular Evolution*, pp. 21-2.

Is a stemma possible for the Hebrew Bible?

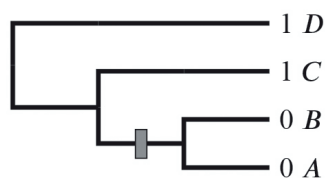


Figure 2: Parsimonious tree.

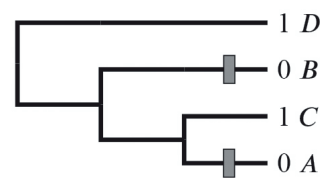


Figure 3: Not parsimonious tree.

In keeping with the principle of maximum parsimony, therefore, the best evolutionary hypothesis is the one that reduces homoplasia to a minimum, producing the shortest phylogenetic tree, that is, the one which contains the least possible number of changes to explain the initial data.

In brief, to identify the most parsimonious tree, the algorithm must perform two tasks: (1) the production of the list of all possible trees; and (2) the calculation of the length of each tree. In the end, the least long tree is selected as the best.

Simply put, to calculate the length of each tree the algorithm assigns to each intermediate node – the common ancestor – the variants present in its descendant members. For a detailed description of the complex assignment process, we refer the reader to the specialist bibliography.⁵⁶ Suffice it to say here that the assignment of such variants, called *ancestral*, has as its primary objective the minimization of the development of independent changes, as already seen. Roughly speaking, once the hypothetical variants of each ancestor have been reconstructed, the algorithm simply counts how many times a given variant changes from one node to another. The final ‘score’ represents the length of the tree.

This process is repeated for all the trees in the list. In theory, the complete list of trees is compiled by exhausting all the possible combinations of witnesses. This can be computationally

prohibitive, however, if the number of witnesses is high, as in our case. A strategy that allows one to ‘get around’ the problem is to launch a so-called heuristic search: in essence, instead of exploring the entire space of possible trees, a heuristic search selects an initial tree and modifies it progressively to minimize its length.

The final step is the selection of the shortest tree. This operation is obvious when the heuristic search results in a single tree. When there are several trees with the same length, one generally proceeds by calculating the so-called strict consensus tree, which results from the sum of only those patterns common to all equally parsimonious trees which have been generated.

The tree shown in Fig. 4 is the result of the consensus of more than 25,000 equally parsimonious trees.⁵⁷ For simplicity, we have placed as the first *taxon* of the tree – to be read from left to right – the text of the reference edition (H), in order to make it easier to understand the position of all the other witnesses with respect to the TR.

Before moving on to describe the consensus tree and to hazard a general interpretation, let us make a few clarifying remarks on the limitations – but also on the usefulness – of the reconstruction we are about to present.

First of all, it is important to reiterate that our tree is not a stemma, for at least two fundamental reasons: (1) it does not display any type of codicological information, such as the date of the witnesses, nor is such information used

⁵⁶ A concise description with further bibliography can be found in M. SPENCER *et al.*, *The Greek Vorlage of the Syra Harclensis: A Comparative Study on Method in Exploring Textual Genealogy*, «TC: A Journal of Biblical Textual Criticism» 7 (2002), <http://rosetta.reltech.org/TC/v07/SWH2002/index.html> (last accessed: 12.6.2021).

⁵⁷ The procedure we have described is identical to that used for the study of the tradition of Grego-

ry of Nazianzus in A.C. LANTIN *et al.*, *Phylogenetic Analysis of Gregory of Nazianzus’ Homily 27*, in G. PURNELLE *et al.* (eds.), *Le poids des mots. Actes des 7èmes Journées Internationales d’Analyse statistique des Données Textuelles*, II, Presses universitaires de Louvain, Louvain-la-Neuve 2004, pp. 700-7. The software used to generate the tree is PAUP, see <https://paup.phylosolutions.com/>.

to determine the tree's structure; (2) it has no direction and therefore does not show the evolution of the text in a temporal sense, from a presupposed ancestor of the whole tradition to its current descendants.

When compared with a stemma, in fact, its limits are indisputable: it can be used neither to represent the *diachronic* evolution of the text (*historia textus*), nor to distinguish among rival variants and therefore offer a tool for the reconstruction of the original (*critica textus*). In short, it is unable to fulfill either of the fundamental purposes which a stemma, and more generally textual criticism, should ideally serve.⁵⁸

In order to use a phylogenetic tree as a stemma, as we have said, it is necessary at least to orient (or root) it;⁵⁹ this operation requires the identification of variants (or MSS) judged to be ancestral, that is, original, and thus presupposes a comprehensive, *qualitative*, study of the whole textual tradition, in other words, the var-

iants of the ancient Versions to which we do not have access at this stage.

Here, however, these limitations and shortcomings are of no concern, since we do not set ourselves as a goal either to reconstruct the history of the tradition (see Sacchi, § 4) or to engage in textual criticism (Borbone). As we specified both in the Introduction and previously in the present section, our analysis is much more limited: what we have in mind is to validate the possibility of organizing medieval documentation into groups or families, to evaluate the quality of the groupings obtained, and to try to outline, through these, a general, *synchronic*, picture of the transmission of Q in the Middle Ages. For this purpose – which, though limited, must precede any future work on the history of the tradition or on textual criticism as traditionally understood – the calculation of a consensus tree is more than sufficient.

⁵⁸ See CHIESA, *Textual History*, p. 264, and similarly R. HENDEL, *The Epistemology of Textual Criticism*, in A.B. PERRIN *et al.* (eds.), *Reading the Bible in Ancient Traditions and Modern Editions*, Society

of Biblical Literature, Atlanta 2017, pp. 252ff.

⁵⁹ On orientation procedures, see SALEMANS, *Building Stemmas*, pp. 31-5.

Is a stemma possible for the Hebrew Bible?

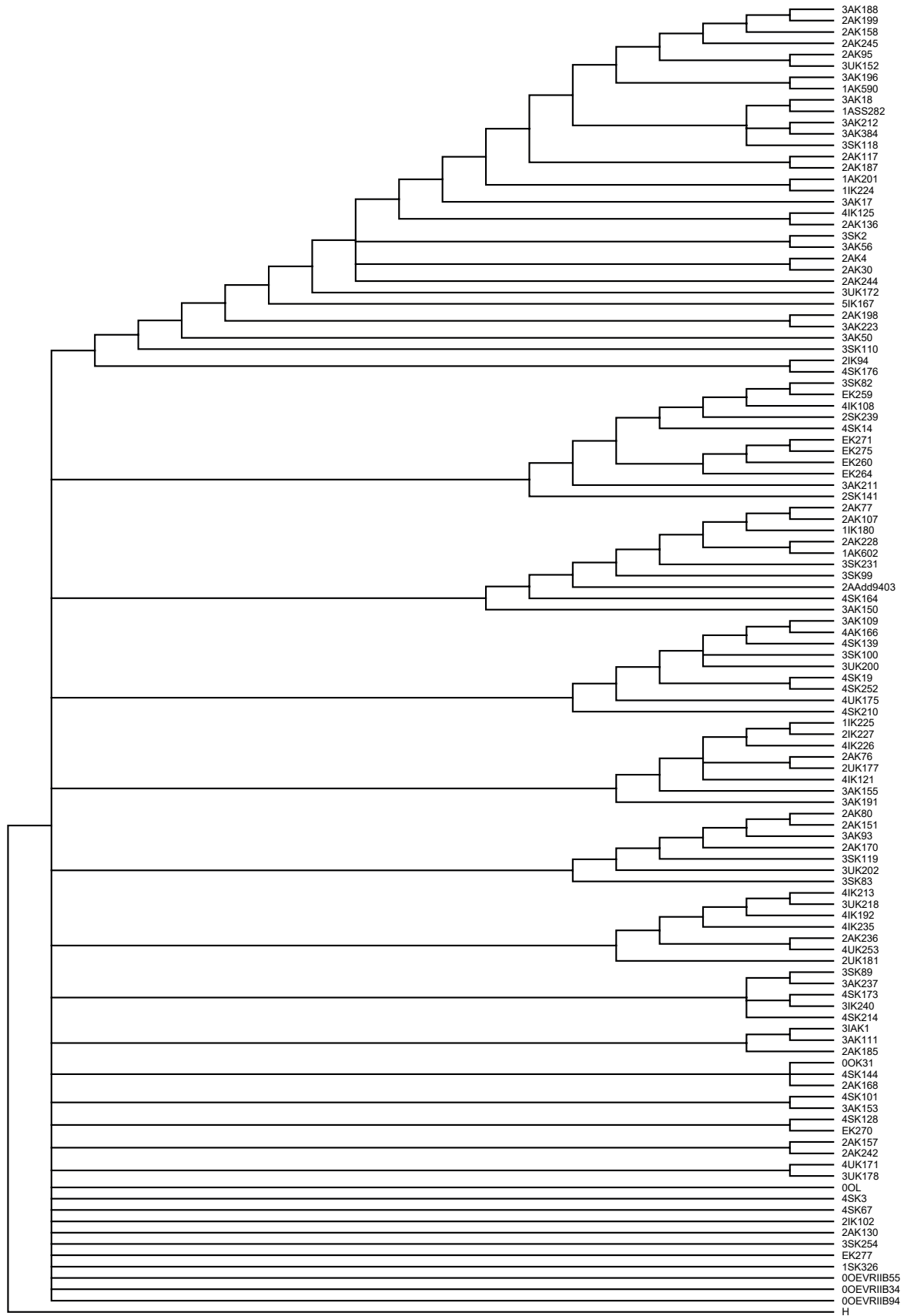


Figure 4: Strict consensus tree of the medieval manuscripts of Qohelet.

5 Results

Our phylogenetic analysis produced 14 groups, each headed by a hyparchetype. Each hyparchetype is assigned a certain number of ancestral variants, which are inherited by all or most of the descendant members. In total, 104 out of 371 variant locations are identified by the algorithm as containing ancestral variants. This means that the variants from other locations are treated by phylogenetic analysis as innovations introduced by individual witnesses. The list of these variant locations, each with a numerical identifier, can be found in Appendix B.

The tree has a typical dichotomous or binary structure, as can be seen: witnesses are united to their respective ancestor in pairs, all the way up to the common ancestor of the entire group. When this does not happen, and a polytomous or multifurcating group is displayed, then the algorithm was not able to find sufficient discriminating information.

Out of 116 witnesses, 105 were allocated into groups: this means that all of the 25,000 equally parsimonious trees which were generated agree on the position to be assigned to the great majority of the witnesses examined. Only in 11 cases was it not possible to find a specific location. These concern those nodes, positioned on the left, which descend directly from the upper horizontal line that constitutes the structure of the tree. However, one need only glance at the table in Appendix B to realize that these are witnesses with very few substantial variants. Here we find the second Rabbinic Bible of Ben Chayyim (EK277), four Sephardi, and four ancient Oriental MSS. Among these, K67, which has no substantial variants, stands out in particular. That these MSS belong to the TR is therefore a certainty.

The groups vary in size and have been sorted progressively, from left to right, on that basis. Below, the reader will find the list, in which each group is accompanied by a brief description indicating: the *sigla* of the witnesses that compose it, the total number of witnesses, and their common ancestral variants. The total number of variants is always provided; as to the numerical identifier, we show it only if the variants are few: in this way the groupings can be qualitatively verified by consulting the table in Appendix B; if there are many variants, on the other hand, their sheer quantity will be sufficient to guarantee the validity of the group.

Finally, we will attempt to call attention to the so-called characteristic ancestral variants, that is, variants which group together two or more witnesses. We call these variants *exclusive*, if they are present only in the witnesses of a single group, and *characterizing*, if present in that same group plus a few additional groups or witnesses.

Here is the list of the 14 groups with commentary:

- 1) Group 4UK171-3UK178; total: 2; variants: 4 (187, 197, 312, 320).

Unclassified MSS with about ten variants each with respect to the TR. The ancestral variants are present in both MSS. Most concern the absence or presence of the copulative conjunction and a case of Ketib-Qere. These variants are also found in a number of other groups.

- 2) Group 2AK157-2AK242; total: 2; variants: 3 (160, 187, 312).

Ashkenazi MSS with about ten variants each with respect to the TR. This group results as weak: the variants are few and polygenetic (variants of the copulative conjunction and an omission due to homeoteleuton).

- 3) Group EK270-4SK128; total: 2; variants: 3 (46, 147, 197).

This group contains the printed edition of the Complutensian Bible and a Sephardi MS. The common variants are weakly kinship-revealing.

- 4) Group 4SK101-3AK153; total: 2; variants: 3 (187, 233, 362).

MSS with about ten variants each. The ancestral variants are weak and not exclusive (one Ketib-Qere and two concerning particles).

- 5) Group 0OK31...2AK168; total: 3; variants: 1 (20).

This is a so-called polytomous group, with multiple branches connected to the same hyparchetype. Whether these MSS belong to a distinct group is questionable (i.e., only one ancestral variant).

- 6) Group 3IAK1...2AK185; total: 3; variants: 4 (79, 102, 187, 312);

- 6a) subgroup 3IAK1-3AK111; variants: + 4 (147, 150, 197, 215).

This is one of the most uniform groups: in fact, it contains four ancestral variants common to all descendants. To these are added another four for subgroup 6a, which therefore has a total of eight shared variants. This group has a characterizing variant (79), attested only in the descendants of this group and in two other MSS with many variants (K158 and K218).

- 7) Group 3SK89...4SK214; total: 5; variants: 2 (23, 197);
 7a) subgroup 3SK89-3AK237, variants: + 1 (349);
 7b) subgroup 4SK173-3IK240, variants: + 1 (365).

The MSS of this group have on the whole fewer than ten variants each as compared to the TR. The ancestral variants are present in all descendants. No. 365 is characterizing for subgroup 7b (found only here and in K107). The group contains polytomies.

- 8) Group 4IK213...2UK181; total: 7; variants: 4 (150, 187, 258, 335);
 8a) subgroup 4IK213-3UK218; variants: 11 (81, 83, 120, 147, 150, 170, 206, 215, 258, 271, 335);
 8b) subgroup 2AK236-4KU253; variants: 6 (71, 83, 120, 150, 187, 258).

MSS mostly Italian and three unclassified. The ancestral variants are not common to all descendants. A characterizing variant is 206, present only in 8a and in the MS with the most variants of any (3UK152).

- 9) Group 2AK80...3SK83; total: 7; variants: 3 (161, 197, 310);
 9a) subgroup 2AK80...3AK93; variants: 9 (83, 102, 120, 152, 161, 167, 233, 310, 312);
 9aa) subgroup 2AK80-2AK151 variants: + 4 (197, 218, 276, 320).

MSS including four Ashkenazi with many variants as compared to the TR and two Sephardi. Ancestral variants are not present in all descendants. The most noteworthy subgroup is 9a (9 variants), and in particular K80-K151 with 12 variants of which one is characterizing (218).

- 10) Group 1IK225...3AK191; total: 8; variants: 2 (215, 312);
 10a) subgroup 1IK225...4IK226; variants: + 4 (6, 223, 233, 258);
 10aa) subgroup 1IK225-2IK227; variants: + 3 (20, 124, 320);
 10b) subgroup 2AK76-2UK177; variants: + 8 (233, 320, 150, 271, 73, 102, 197, 276).

MSS mostly Italian. This is one of the groups exhibiting the most consistency, with ancestral variants inherited by all members. Within this group, note K76-K177 (10 variants) and the Italian subgroup K225...K227 (6), especially K225-K227 (9). MSS K225-K226 have two characterizing variants (in increasing order of occurrence in the other branches of the tree, 45 and 30). Among the common variants, the pair

K76-K177 includes an exclusive variant (73).

- 11) Group 3AK109...4SK210; total: 9; variants: 2 (271, 310);
 11a) subgroup 3AK109...4SK139; variants: + 5 (20, 71, 83, 150, 312);
 11aa) subgroup 3AK109-4AK166; variants: + 5 (167, 187, 314, 335, 364);
 11b) subgroup 4SK19-4SK252; variants: + 3 (120, 197, 292).

MSS mostly Sephardi and two Ashkenazi; the latter have a large number of variants with respect to the TR, while the Sephardi have about ten variants each. Only the first of the ancestral variants is common to all descendants. The most significant subgroup of this unit is 11a (7 variants), and K109-K166 in particular (12). This pair has a characterizing variant (364), present here and in four other MSS. The subgroup of two Sephardi K19-K252 might also be noted, with five common variants.

- 12) Group 2AK77...3AK150; totals: 10; variants: 2 (187, 215).
 12a) subgroup 2AK77...1IK180; variants: 8 (102, 20, 23, 310, 120, 320, 152, 295);
 12aa) subgroup 2AK77-2AK107; variants: + 11 (12, 51, 112, 150, 223, 233, 261, 292, 312, 337, 349);
 12b) subgroup 2AK228-1AK602; variants: + 11 (102, 271, 20, 23, 310, 312, 98, 120, 223, 320, 50).

MSS including six Ashkenazi, three Sephardi, and one Italian. Common variants are not inherited by all members. In particular, the Italian-Ashkenazi group made up of K77...K602 (4 variants) stands out, which in turn can be divided into two groups with two hyparchetypes: K77...K180 (8) and K228-K602 (13). MSS K77 and K107 are among those with the greatest number of variants in common (19), including one which is exclusive (112).

- 13) Group 3SK82...2SK141; total: 11; variants: 3 (187, 271, 312).

This group is characterized by being in large part made up of printed editions. It is among the groups which display the greatest cohesion, with ancestral variants distributed to all descendants. There are two main groupings: one containing the edition of the Naples Hagiographa (EK259) and the other consisting of two editions of the First Rabbinic Bible (K271, K275), the Soncino Bible (K260), and the Brescia Bible (K264). The first, according to the tree, is close to MSS K82 and K108 (8 common vari-

ants), and especially to the first of these (14). In the second, subgroup K271-K275 (9) stands out in particular, with an exclusive variant (353).

14) Group 2AK188...4SK176; total: 33; variants 2 (197, 215)

Group with a clear Ashkenazi majority. It contains MSS with many substantial variants with respect to the TR. The second ancestral variant is found in all but three descendants. Within this group we find numerous subgroups consisting of two MSS. Among these we might mention, in order of appearance in the tree: K188-K199 (15 variants), K95-K152 (27), K196-K590 (15), K18-SS282 (26), K212-K384 (27), K117-K187 (20), K201-K224 (22), K125-K136 (20), K2-K56 (10), K4-K30 (17), and K198-K223 (8). Among the variants in common, the K4-K30 pair has two exclusive variants (176, 179) and a characterizing one (246). Another group with exclusive variants is K95-K152 (54, 100, 185).

6 Discussion

In our Introduction we set for ourselves a twofold challenge: to verify whether witnesses could be grouped using philological criteria, and to describe the physiognomy of the Hebrew medieval textual tradition based on the groupings so obtained.

With regard to the first objective, the method applied has indeed produced groups, demarcated by the distribution of ancestral variants. Some of these, as we have seen, rest on fragile foundations, in the sense that either the number or type of inherited variants is insufficient to affirm with certainty a genetic relationship. This is the case with groups 1-5 and 7, mainly composed of MSS which differ little from the TR. No

characteristic variants could be found for these. For other groups, however, the hypothesis of a common ancestor is likely and calls for further investigation. Among those that seem most probable are groups 6, 10, and 12-14. Generally speaking, these are composed mostly of Italian and Ashkenazi witnesses with many common substantial variants, including a certain number of characteristic variants. One of the largest and most uniform subgroups is that of the Italian codices K225, K226, and K227 (group 10). The other subgroups mostly consist of pairs of MSS. In terms of quantity, those in group 14, the largest revealed by the machine, stand out above all.

Such a division makes it possible to speculate on the history of the manuscript tradition of Q. Most of the codices whose classification was either impossible or doubtful can be traced back to a branch of the tradition very close to the TR. The few attested variants – onto which the algorithm has occasionally managed to structure groups – can be interpreted as isolated variants of the TR.

The codices that have instead been assigned to larger groups, above all group 14, seem to belong to a separate branch of the tradition, which could be identified with that generally known as *non-receptus* or *anti-receptus* (AR).⁶⁰

In medieval Europe, starting from about the year 1000, two main branches of the HB tradition seem therefore to develop in contraposition to each other: one of Sephardi origin, inspired by the Ben Asher Oriental recension, the other, which, following Sacchi, we might call Western,⁶¹ typically Italian and Ashkenazi.⁶²

The existence of two branches of the tradition is corroborated by the fact that the variants are mostly two in number: of the 371 variant locations examined, it turns out that only 20 contain more than two readings: the vulgate reading

⁶⁰ In the past, the definition ‘amasoretic’ was used for these codices, see DE ROSSI, *Compendio di critica sacra. Dei difetti e delle emendazioni del Sacro testo e piano d’una nuova edizione*, Parma 1811, pp. 9ff., and L. BERTHOLDT, *Historischkritische Einleitung in sämtliche kanonische und apokryphische Schriften des alten und neuen Testaments*, II, Erlangen 1813, pp. 437ff. Today the prevalent terms are *anti-receptus* (Sacchi) or *non-receptus* (GOSHEN-GOTTSTEIN, *The Rise of the Tiberian Bible Text*, in A. ALTMANN (ed.), *Biblical and Other Stud-*

ies, Harvard University, Cambridge (MA) 1963, pp. 108ff).

⁶¹ SACCHI, *Analisi quantitativa*, p. 12.

⁶² The study of the medieval biblical tradition along ethno-geographic lines was conducted above all by J. Penkower in his fundamental work, unfortunately not published and difficult to find, cited in note 12: PENKOWER, יעקב בן חיים. A summary can be found in ID., *Rabbinic Bible*, in J.H. HAYES (ed.), *Dictionary of Biblical Interpretation*, II, pp. 361-4, Abingdon, Nashville 1999, and in ID., *The Devel-*

and its variant. The oscillation is therefore almost always between TR and AR.⁶³

The same consideration applies to scribal corrections. As previously stated, these concern two types: original variants corrected in the TR sense (*primo*) and readings of the TR corrected in the AR sense (*nunc*). Cases of duplicate variants resulting from correction are very rare.⁶⁴

Corrections of the first type, on the other hand, are very frequent and are testimony to the great influence of the recension of the TR, especially in the Ashkenazi area.⁶⁵ Limiting ourselves to the MSS we were able to re-examine, we might mention 3AK17 (36 *primo* variants out of 51), 1ASS282 (33/62), 3UK218 (29/41), 2AK136 (25/33), 3AK212 (25/44). Even some Sephardi MSS have a certain number of corrections in accordance with the TR, at times enough to cover the entire array of variants examined, such as for 3SK254 (6/6) and 3SK82 (22/25).

Nevertheless, the authority of the TR has not been able to completely extinguish the reading tradition of the AR. In fact, as Sacchi asserts, “in certain places one could also read the A.R. reading with such conviction as to correct the ms

in the A.R. sense.”⁶⁶ Corrections in the AR sense are rarer, but they do exist, and these, too, particularly in the Italian-Ashkenazi area. The MSS that present multiple such variants are 1AK201 (9), 3AK212 (7), and 1IK180 (6), the first two belonging to group 14, the third to group 12. Several such variants have survived both in the text of printed editions as well as in the form of *marginalia*.

With respect to the antiquity of AR variants, as we anticipated in our Introduction, the lack of data relating to the ancient Versions does not allow us to specify whether they are pre-Masoretic or not. If they are, then the AR variants may not, at least not always, be the fruit of scribal errors, but rather the remnants of a very ancient tradition;⁶⁷ if they are not, then such variants presumably arose later, indeed in medieval times. It is clear that the adoption of one or the other of the two models has significant consequences for textual criticism: if we hold with the first, then the AR MSS would be crucial, in that they would retain “extremely ancient variants worthy of the utmost consideration”;⁶⁸ if, instead, we favor the second, then the AR MSS would find themselves

opment of the Masoretic Bible, in A. BERLIN - M.Z. BRETTLER (eds.), *The Jewish Study Bible*, pp. 2077-84, Oxford University, Oxford/New York 2014.

⁶³ Also noted by SACCHI, *Analisi quantitativa*, p. 11.

⁶⁴ One may cite the example of 3SK110 in Q 9,2: the TR reads ולטהור, the original reading of the MS is וטהור, subsequently corrected in לטהור also attested in other MSS.

⁶⁵ The authority of the TR in general and of the “accurate Sephardi MSS” (“ספרי ספרד המדוייקים”) in particular was already recognised in the Middle Ages, especially in the Ashkenazi area, see M. COHEN, האידדיאה בדבר קדושת הנוסח.

⁶⁶ SACCHI, *Analisi quantitativa*, p. 12.

⁶⁷ This is the thesis supported by Sacchi and accepted by Chiesa, with the fundamental argument that almost none of the variants of the Hebrew MSS is isolated, that is, not supported by at least one of the ancient Versions, see SACCHI, *Analisi quantitativa*, pp. 10ff., and CHIESA, *L'Antico Testamento*, pp. 279-80. This thesis has been criticized by Barthélemy as vitiated by the selection criteria of DR mentioned in our Introduction, see BARTHÉLEMY, *Critique textuelle*, pp. xx-xxv. Among the studies intended to investigate the relationships between

medieval MSS and ancient Versions (see note 12) we recall that of Wevers, according to whom the existence of pre-Masoretic variants is an undeniable fact that undermines the theory of a single source (§ 2), see WEVERS, *A Study*, p. 76. A similar position is expressed in J. TREBOLLE BARRERA, *The Jewish Bible and the Christian Bible. An Introduction to the History of the Bible*, Brill, Leiden 1998, pp. 280-1. Even scholars who generally agree in denying authority to medieval MSS admit that their variants can sometimes be considered original. In the words of Goshen-Gottstein: “[a]lmost all our evidence from medieval MSS would be explicable as a secondary development from a common archetype and practically all of it as belonging to one ‘recension’. Were it not for the disturbing ‘almost’, the whole chapter on medieval MSS could be regarded as closed and our apparatus be freed from them once and for all,” GOSHEN-GOTTSTEIN, *Hebrew Biblical Manuscripts*, pp. 285-6. So also Barthélemy, who studied the variants of several of these MSS, such as K150, K93 and K96. According to him, a minority of cases (28 out of 334) could attest to ancient variants independent of the Tiberian tradition, see BARTHÉLEMY, *Critique textuelle*, pp. xxxii-xlix.

⁶⁸ SACCHI, *Analisi quantitativa*, p. 12.

in a subordinate position within a hypothetical stemma, and their variants would tend to be excluded as being secondary innovations.⁶⁹

What we can confidently affirm here is that some of the codices classifiable as AR are indeed relatively ancient. A certain AR-type reading tradition of Q is therefore well rooted in the Europe of the Italian-Ashkenazi area, at least from the 12th century.

7 Conclusions and perspectives

Our method has permitted us to meet the twofold challenge we set for ourselves at the outset: to group textual witnesses and to outline the medieval Hebrew tradition of Qohelet. Some of the groupings, as we have seen, are probably artificially created by the machine and do not withstand a qualitative analysis. Others are quite likely to be authentic, as in the case of the printed editions and of the Italian and Ashkenazi codices. The theory of a bipartition between *textus receptus* and *anti-receptus*, formulated in other studies of textual history, fits well with the results we have obtained and may in effect be considered as corroborated by them. It goes without saying that a qualitative analysis, not only textual but also codicological and paleographical, is essential here, for only in this way may we reach a definitive conclusion as to the true efficacy of the method.

The limitations of our reconstruction, as we have underlined more than once, are evident: For one thing, the significant variants on which we based our analysis are few, and not all of them

possess the same ability to reveal genealogical relationships. Moreover, contamination, which is more than likely in a case like the medieval tradition of the Hebrew Bible, is in fact ignored by us. And finally, the sample we analysed is extremely small – barely more than a hundred witnesses, as against a tradition that contains thousands.⁷⁰ To this must naturally be added collation errors, both our own and those of Kennicott.

To try to overcome these problems, “it is necessary”, as Borbone so justly affirms, “to proceed beyond the collations of Kennicott and De Rossi.”⁷¹ An entirely new collation of medieval manuscripts, conforming to modern philological criteria, is crucial to arrive at a delineation of a *stemma codicum* of the medieval Hebrew biblical tradition. Such a collation must include new data, as Barthélemy suggests, such as the variants of vocalization and punctuation, and the Massora.⁷² Other variants that might turn out to prove genealogically informative – and that are not documented in the traditional collations – are para-textual variants, such as those concerning division of the text into units. Finally, it will be important to include other manuscripts, whether older or relevant from a text-critical point of view, if we want to obtain a less fragmented picture of the transmission of the Hebrew biblical text in the Middle Ages.

Having said all of this, it does seem to us appropriate to emphasize in conclusion that the problems referred to – hegemony of the *textus receptus*, contamination, overabundance of witnesses – are not exclusive to the Old Testament tradition. New Testament scholars find themselves in a similar situation. There, however, the

⁶⁹ For Goshen-Gottstein, even if we admit that in some cases the MSS may retain ancient variants, this model remains the best solution from a practical point of view: in fact, “the possibility of infiltration of extra-Massoretical variants into the Massoretical ‘central current’ is so negligible that for all practical purposes it may be disregarded [...] the medieval Hebrew MSS are therefore without practical value for any attempt to reach back into the early history of the Bible text,” GOSHEN-GOTTSTEIN, *Hebrew Biblical Manuscripts*, p. 286. According to the author, these variants do not reflect an independent Hebrew text type, but are rather “variant-splinters, odd freaks of individual survivals,” M.H. GOSHEN-GOTTSTEIN, *The Development of the Hebrew Text of the Bible:*

Theories and Practice of Textual Criticism, «Vetus Testamentum» 42 (1992), p. 209.

⁷⁰ If to these limitations we add the number of lost witnesses, it can be assumed that there is little hope of tracing the *descripti*. As Brambilla Ageno writes for the romance texts: “[i]t should be kept in mind that in general, given the huge mass of lost codices (certainly more numerous than those preserved), the probability that an exemplar and its copy remain is minimal,” F. BRAMBILLA AGENO, *L’edizione critica dei testi volgari*, Antenore, Padova 1999², p. 98.

⁷¹ BORBONE, *Osea*, p. 193.

⁷² BARTHÉLEMY, *Critique textuelle*, p. xxvii.

methodological reflection on the applicability of the genealogical model is central and long-standing,⁷³ and has led to important developments, having to do specifically with the problem of contamination and, more generally, with the analysis of open traditions.⁷⁴ Other traditions, these, too, contaminated and characterized by a strong tendency towards standardization to a vulgate, have also been treated with computational stemmatology techniques similar to those used by us, with most interesting results.⁷⁵

Here, we have sought to explore the potential of such techniques, by applying to the medieval tradition of the Hebrew Bible the same genealogical method practiced by New Testament and other philologists. This is an undertaking which, as far as we are aware, has never previously been attempted, and the presumed unfeasibility of which has in our opinion been too hastily – if understandably – asserted.

For our part, we intend this article to represent an initial step forward, not so much to an-

swer as to properly frame the question we have posed in our title, spurred on by the twofold conviction that stemmatic analysis is an avenue that can and must be pursued, and that neither the peculiarities presented by any particular case study, nor the limits inherent in the investigative tools currently available, should be invoked to assert *a priori* the impossibility of working with classic philological methods.

Appendix A: List of witnesses

Below we present the list of witnesses with the total of the variants by category. The re-collected MSS are indicated by an asterisk. For completeness, we list the total of all variants which are present in KN, while in the last column we report the total of the variants we actually analysed. The list is sorted in descending order according to this last.

Witness	Accidentals	Substantials	<i>Singulares</i>	<i>primo</i>	<i>nunc</i>	total	#
3-UK152	242	162	144	5	1	404	76
2-AK77	171	106	60	1	0	277	64
3-AK384	109	94	26	1	1	203	63
1-ASS282*	17	97	32	71	2	114	62
2-AK95*	148	105	48	20	0	253	59
3-AK18	133	88	35	2	0	221	55
3-AK17*	120	91	40	79	1	211	51
2-AK107*	119	116	74	77	0	235	50

⁷³ For example, the *Claremont Profile Method*, see B.M. METZGER - B.D. EHRMAN, *The Text of New Testament. Its Transmission, Corruption, and Restoration*, Oxford University, New York/Oxford 2005⁴, pp. 231-49. See also J.R. ADAIR, *Old and New in Textual Criticism: Similarities, Differences, and Prospects for Cooperation*, «TC: A Journal of Biblical Textual Criticism» 1 (1996), <http://jbtc.org/v01/Adair1996.html> (last accessed: 12.6.2021).

⁷⁴ This is the case with the *Coherence-Based Genealogical Method*, see G. MINK, *Problems of a Highly Contaminated Tradition: The New Testament*, in P. VAN REENEN *et al.* (eds.), *Studies in Stemmatology*, II, pp. 13-85, and P.J. GURRY, *A Critical Examination of the Coherence-Based Genealogical Method in New Testament Textual Criticism*, Brill,

Leiden/Boston 2017.

⁷⁵ These are the studies on the tradition of Gregory of Nazianzus (see also note 57): P.V. BARET *et al.*, *Experimental Phylogenetic Analysis of a Greek Manuscript Tradition*, «Journal of the Washington Academy of Sciences» 89 (2003), pp. 117-24; C. MACÉ *et al.*, *Le classement des manuscrits par la statistique et la phylogénétique: le cas de Grégoire de Nazianze et de Basile le Minime*, «Revue d'Histoire des Textes» 31 (2001), pp. 241-73; C. MACÉ, *Gregory of Nazianzus' Homilies. An Over-Abundant Manuscript Tradition in Greek and in Translation*, in A. BAUSI *et al.* (eds.), *Comparative Oriental Manuscript Studies. An Introduction*, Tredition, Hamburg 2015, pp. 424-9.

Luigi Bambaci

Witness	Accidentals	Substantials	Singulares	primo	nunc	total	#
2-AK199	16	74	26	5	0	90	48
3-AK212*	119	55	9	82	7	174	44
2-AK151	91	65	21	17	1	156	43
2-AK80	87	85	42	5	0	172	43
2-AK117*	134	56	14	12	0	190	43
2-AK4*	76	65	23	33	3	141	42
2-AK30	106	57	18	0	0	163	42
2-AK76	81	81	38	2	0	162	42
3-UK218*	104	59	16	50	1	163	41
1-IK224	131	44	7	0	0	175	39
2-AK187	81	60	26	0	0	141	38
4-IK125*	79	47	9	8	0	126	38
1-AK201*	17	56	8	30	9	73	37
1-IK180*	15	65	22	31	6	80	37
3-SK118*	84	50	16	12	0	134	34
3-AK196	94	48	16	0	0	142	33
2-AK136*	98	61	30	58	0	159	33
1-AK590*	16	43	10	16	0	59	32
5-IK167*	64	97	66	6	0	161	31
3-AK56*	101	55	26	13	0	156	30
3-AK109*	101	60	31	21	0	161	30
2-AK245*	89	44	12	29	0	133	30
1-AK602*	15	35	3	7	3	50	29
EK259	80	35	8	0	0	115	28
3-AK188	91	29	1	1	0	120	28
2-AK228*	18	32	4	6	2	50	28
4-AK166*	71	42	12	14	0	113	28
3-IAK1	46	47	23	3	1	93	27
3-AK111*	77	58	30	34	1	135	27
2-AK158*	68	35	6	28	3	103	26
3-SK82*	76	32	9	45	1	108	25
3-AK211*	67	43	21	35	0	110	25
3-SK2*	83	37	15	8	0	120	25
3-AK93	90	37	17	2	0	127	25
3-AK223	49	43	17	0	1	92	25
2-UK177*	87	36	10	17	2	123	24
4-IK213*	113	33	11	25	1	146	23
4-IK226*	56	38	17	14	0	94	22
4-IK121*	96	40	24	2	1	136	22
2-IK227*	32	38	16	13	0	70	22
2-AAAdd9403*	10	36	14	27	1	46	22

Is a stemma possible for the Hebrew Bible?

Witness	Accidentals	Substantials	<i>Singulares</i>	<i>primo</i>	<i>nunc</i>	total	#
4-SK14*	101	33	8	15	2	134	22
2-AK244*	116	34	15	30	1	150	22
3-SK99*	91	37	14	13	4	128	21
1-IK225*	26	23	4	5	0	49	19
2-AK170*	68	42	25	26	2	110	19
3-AK150*	54	36	18	22	1	90	18
2-SK239*	61	20	2	1	0	81	18
2-AK198	5	24	6	1	0	29	18
3-UK202	4	19	2	1	1	23	16
3-UK172	94	17	1	4	1	111	16
3-SK119*	60	20	4	7	1	80	15
4-IK108*	77	21	3	0	3	98	15
3-UK200	12	23	7	2	0	35	15
4-SK19*	77	28	14	7	0	105	14
3-AK153	50	23	10	0	0	73	14
3-AK155*	45	18	4	7	0	63	14
4-UK253	109	18	4	0	0	127	14
4-SK252	88	21	9	1	0	109	14
3-SK231	34	13	0	1	0	47	13
4-IK235*	38	20	6	9	1	58	13
4-IK192	45	18	5	0	0	63	13
EK271	24	12	0	0	0	36	12
4-UK171*	8	14	3	9	1	22	12
EK264	27	13	1	0	0	40	12
2-AK236*	43	19	6	10	1	62	12
2-AK157*	27	42	28	29	4	69	11
3-UK178	53	17	7	0	1	70	11
3-SK100*	20	14	3	4	1	34	11
2-AK242*	12	14	1	9	4	26	10
4-SK173	64	13	4	0	1	77	10
3-AK191	51	12	2	0	0	63	10
4-SK128	46	15	4	24	1	61	10
4-SK101*	44	20	11	1	0	64	9
3-IK240	94	17	8	0	0	111	9
EK260	75	12	3	0	0	87	9
3-AK50	18	11	2	1	0	29	9
EK275	28	11	2	0	0	39	9
4-SK139	25	8	1	15	0	33	8
2-UK181	36	18	8	0	3	54	8
4-SK210*	19	11	3	1	0	30	8
3-AK237	65	10	2	2	0	75	8

Luigi Bambaci

Witness	Accidentals	Substantials	<i>Singulares</i>	<i>primo</i>	<i>nunc</i>	total	#
4-SK164*	63	11	5	5	0	74	7
4-SK3*	51	10	3	6	0	61	7
4-UK175	28	9	2	1	0	37	7
2-AK185	13	8	1	0	0	21	7
2-AK168*	39	14	9	12	2	53	6
2-SK141	34	8	0	15	2	42	6
4-SK176	61	6	0	0	0	67	6
3-SK254*	13	9	1	8	3	22	6
3-SK83*	19	6	1	2	0	25	5
3-SK110*	40	8	1	3	1	48	5
2-IK94	36	5	1	8	0	41	5
0-OEVRIIB55*	3	4	0	2	0	7	4
2-AK130*	8	9	5	5	1	17	4
EK270	34	5	1	0	0	39	4
3-SK89	33	10	5	1	0	43	4
2-IK102*	37	5	2	5	0	42	4
0-OK31*	22	9	6	4	1	31	4
0-OEVRIIB94*	2	5	2	1	0	7	3
4-SK214*	20	7	3	2	2	27	3
1-SK326*	7	5	2	7	0	12	3
4-SK144*	21	7	4	3	0	28	3
0-OEVRIIB34*	3	2	0	2	0	5	2
0-OL*	20	1	0	0	0	21	1
4-SK67*	20	0	0	0	0	20	0
EK277	3	0	0	0	0	3	0

Is a stemma possible for the Hebrew Bible?

Appendix B: List of ancestral variants

Below is the list of variant locations containing ancestral variants. These variants, we recall (§§ 4,5), are those which the algorithm has assigned to the intermediate nodes of the tree (the hyparchetypes) and which are supposed to be inherited from the terminal nodes (the witnesses). In total, the variant locations chosen by the algorithm are 104 out of 371 (hence the discontinuous numbering). For each of these we

report: the numerical identifier (ID), the Qohelet passage concerned, the lemma of the reference text used by Kennicott, and the variant(s). The number that is sometimes found immediately before the lemma indicates the order of occurrence of the word(s) in the reference text (e.g., 'לא 2' in ID 6 means: the second occurrence of the word 'לא' in Qohelet 1,8). In those cases where there are more variants, these are separated by a vertical line ' | ' (see ID 52). An omission is indicated by a hyphen '-'.

ID	passage	lemma	variant location
2	1,5	ואל	אל
6	1,8	לא 2	ולא
7	1,9	ואין	אין
10	1,11	לא	ולא
12	1,13	השמים	השמש
19	1,16	הנה	-
20	1,16	על ירושלים	בירושלים
21	1,17	לבי	את לבי
23	1,17	ושכלות	וסכלות
30	2,4	נטעתי	ונטעתי
32	2,5	ונטעתי	נטעתי
34	2,7	קניתי	קניתי לי
36	2,7	שהיו	אשר היה
42	2,10	לא 2	ולא
45	2,10	מכל 2	בכל
46	2,12	ושכלות	ושכלות
50	2,14	הולך	ילך
51	2,14	וידעתי	ידעתי
52	2,14	גם	- כי
54	2,15	אני 3	- גם אני
65	2,20	שעמלתי	שעמלתי ושחכמתי
71	2,24	היא	הוא
73	2,26	ודעת	-
79	3,11	לא	-
81	3,13	בכל עמלו	בעמלו
83	3,13	היא	הוא
85	3,15	את	-
98	3,19	ומקרה 2	מקרה
100	3,19	כן מות זה	-
102	3,21	מי	ומי

Luigi Bambaci

ID	passage	lemma	variant location
112	4,4	אני	-
113	4,4	היא	הוא
120	4,8	עיניו	עינו
124	4,14	הסורים	האסורים
127	4,16	לכל 2	-
128	4,16	גם 1	וגם
131	4,16	כי	-
134	5,1	על 1	את
138	5,3	אל	לא
139	5,5	ואל	אל
143	5,6	כי 2	-
147	5,8	היא	הוא
150	5,10	ראית	ראות
152	5,11	ואם	אם
160	5,17-18	כי...האלהים	-
161	5,18	גם	וגם
162	5,18	ונכסים	ונסכים וכבוד
167	5,18	היא	הוא
169	6,1	רעה	רעה חולה
170	6,1	אשר	-
171	6,1	היא	הוא
173	6,2	זה	גם זה
176	6,3	לו	בו
179	6,5	שמש	השמש
183	6,8	מה 2	ומה
185	6,9	מהלך	מהלוך
186	6,10	ולא	לא
187	6,10	שתקיף	שתקיף
191	6,12	חיי	- חיי
194	7,2	משתה	המשתה
196	7,6	כקול	בקול
197	7,6	גם	וגם
200	7,8	רוח 1	אפים
206	7,12	בעליה	את בעליה
215	7,18	את 1	-
218	7,21	לכל	כל
223	7,25	ולבי	- לבי בלבי את לבי
225	7,25	הוללות	והוללות
231	8,1	ישנא	ישונה
233	8,3	אל 2	ואל

Is a stemma possible for the Hebrew Bible?

ID	passage	lemma	variant location
235	8,4	מלך	המלך
236	8,4	שלטון	ושלטון
243	8,7	כאשר	באשר
245	8,9	מעשה	המעשה
246	8,9	נעשה	המעשה
250	8,11	בני	-
258	8,13	אלהים	האלהים
261	8,15	ימי	מספר ימי
269	8,17	לא 2	ולא
271	9,1	אל	את
276	9,2	לטוב	- ולטוב
281	9,4	יבחר	יחובר
292	9,9	אשר נתן...הבלך	-
293	9,9	כל ימי הבלך	-
294	9,9	הבלך 2	חיי הבלך
295	9,9	הוא	היא
306	9,12	כהם	בהם
310	9,15	חכם	וחכם
312	10,1	מכבוד	ומכבוד
314	10,3	כשהסכל	שהסכל כשסכל בשהסכל
316	10,4	תעלה	יעלה
320	10,9	בוקע	ובוקע
327	10,14	לא	ולא
335	10,20	הכנפים	כנפים
337	11,3	ואם 1	אם
339	11,3	ואם 2	אם או
349	11,9	ובמראי	ובמראה
353	12,1	ימי	מי
362	12,6	ירחק	ירתק
363	12,7	על	אל
364	12,8	הקוהלת	קהלת
365	12,8	הכל הבל	הבל הבלים הכל הבל
366	12,9	תקן	ותקן תכן
371	12,14	ואם	אם

Luigi Bambaci

SUMMARY

This article presents the results of an experiment of computational stemmatology on medieval manuscripts of the biblical book of Qohelet on the basis of the data provided by Kennicott's collation. The aim of the experiment is to verify the possibility of treating the medieval tradition of the Hebrew Bible as groups or families of textual witnesses, using phylogenetic analysis, as a first step towards the definition of a *stemma codicum*.

KEYWORDS: Textual history of the Hebrew Bible; Hebrew medieval manuscripts; Computational stemmatology.