?

The Witness Relations in The Nun's Priest's Tale and Related Links

Peter Robinson

Introduction

The Nun's Priest's Tale and links related to it (named L30 L31 L32 in the Canterbury Tales Project numbering) are extant in 51 manuscripts and four print editions dating from before 1500. This section presents our analysis of the textual relationships among these witnesses, based on an analysis of our collation and using the various tools available to us.

All 55 witnesses are present for the Nun's Priest's Tale; two (Ra1 Me) lack more than 50% of the text; four lack over 100 lines (Ld1: 364-end; Ra2 1-173; En2 1-28, 568-end; Fi 100-155, and some 150 other lines missing) while the other 53 are relatively complete. Only three lines (255, 256, 259) are present in all 55 witnesses. Link 30, the link usually known as the Nun's Priest's Headlink or Prologue, is present in 50 of these 55 witnesses: that is, all but Gg Gl Ra1 Ra2 Ra3 (En2 has only a few lines of the link). Link 31, sometimes printed as the Epilogue to the Nun's Priest's Tale, is present in eleven witnesses: the five manuscripts affiliated to the a group (Cn Dd Ds1 En1 Ma); the pair Ad1/En3, the related print editions Cx2 Wy, and two manuscripts usually independent of these, Ch and Ry1. Link 32 appears only in the a manuscript pair Cn Ma and in the pair Ad1 En3, following on in these immediately from Link 32.

1. Phylogenetic analysis of the data

1.1 NEXUS files and PAUP

In order to carry out phylogenetic analysis, all data was converted into the NEXUS file format. The NEXUS file format is widely used by evolutionary biologists to hold data concerning agreements and disagreements among populations (‘taxa’ in evolutionary biology; ‘witnesses’ to us) at precise points (‘characters’ in their terms, variants in ours). The fundamental element in a NEXUS file is a data matrix, in which the agreements and disagreements at each place of variation (‘character’) among the objects surveyed are registered as entries in a series of columns and rows. This example shows the variants on the word ‘thus’ in line of Link 1, in NEXUS file data matrix format:

GL1L1_thus 00200100000002000003000002002000?0200002?00000300000022000

The label ‘GL1L1_thus’ line declares that these are the variants at the word ‘thus’ of line 1 of Link 1. Following this label is a series of characters (0 1 2 3 ?). Each place in this series corresponds to a witness. The first place is for the collation base, is always zero, and is later discarded. The second place represents witness Ad1, then Ad2, Ad3, Bo1, Bo2 etc. Ad1 Ad3 and Bo1 all here have reading 0: from another part of the file, we find this is ‘thus’. Bo2 has reading 1: ‘þis’, Ad2 has reading 2, which is omission of this word.

Presenting the data in this form makes it very tractable for computer analysis. We have, after much experiment, settled on the program PAUP (Phylogenetic Analysis Using Parsimony, Swofford 1996) for analysis, though we use other programs for particular situations. Generally, PAUP gives good results where it appears that the population surveyed has developed through comparatively straight-forward genetic descent; that is, in situations where most variation has been propagated through inheritance, from parent to child, and where there has not been large-scale sideways transmission of variation as might occur in a heavily contaminated tradition. PAUP seems particularly fitting for our purposes because of its use of sophisticated methods to find the most ‘parsimonious’ evolutionary hypothesis. Briefly, this method seeks to explain the sharing of characteristics as evidence of common descent, rather than by independent introduction in each object. That is: if reading A is present in two manuscripts, parsimony analysis seeks to explain this by supposing there was one change only in a joint ancestor

of the two, rather than two distinct changes, one in each manuscript. Hence the term ‘parsimony’: this method looks for the genetic hypothesis which requires the least number of changes to explain the distribution of agreements and disagreements in the objects surveyed; that is, the most parsimonious explanation.

In even a moderate size tradition, the number of possible tree arrangements rapidly climbs to huge numbers: so much that to examine every possible arrangement to find out which is the most parsimonious would require more computers and time than any of us will ever have. As a result, PAUP (like other parsimony tools) uses a range of techniques to optimize the search process by avoiding the need to examine every possible tree. In PAUP’s implementation each parsimony replication fixes on an initial tree within an ‘island’ of trees among the many trillion (literally) possible for this set of data. It then examines the trees in this island by branch-bisection and swapping on this initial tree, and on further trees generated from this initial tree. It repeats this process as many times as you ask, and at the end of the process offers the tree or trees with least number of changes (the ‘shortest’, or most parsimonious tree or trees) it has found. It appears that in cases where there is considerable ambiguity in the data (as is likely to be the case where the tradition is heavily contaminated, or there is insufficient data to give a clear analysis) parsimony analysis might yield many trees of identical length at each replication, each of which has to be subjected to branch-bisection and exchange. In these instances, parsimony analyis may be of limited use. In our work with the Canterbury Tales project, we have found this to be the case particularly with smaller ranges of data, that is less than 500 characters (corresponding to about 50 lines of text). It appears that the prevalence of lateral movement of variants (that is, by contamination, shift of exemplars and the like) within the tradition means that over smaller datasamples, a few instances of lateral movement can render the analysis unstable and unreliable. Over larger sets of data, the underlying fundamental groups, corresponding to genetic descent by regular copying of a series of single exemplars into single copies, show through. It is the same principle by which in small population samples, one may encounter many more males than females, but the larger the sample, the more nearly the numbers balance.

PAUP offers a range of methods besides parsimony, and we have found the neighbour-joining method also useful. Instead of a genetic evaluation of the data by hypothesizing modification and descent at each point, it works by evaluating a distance matrix of the total sums of agreements and disagreements across the materials. This procedure identifies ‘neighbours’ among the whole population; these neighbours are then progressively joined together, according to their place in the matrix, to build a branching, tree-like representation.

Both forms of analysis begin by joining the taxa (witnesses, in our terms), and by joining the taxa in different orders one can create differing trees. In a 'perfect' genetic situation, that is one where all differences have arisen by descent with modification and hence with no lateral transmission of variation, one would expect that the order in which the taxa are joined would have little or no effect on the trees found. Accordingly, one can use the degree to which different initial ordering creates different final trees as a measure (though inexact) of the lateral transmission within the tradiion. More particularly, one can use this to determine the relative confidence we have in different areas of the tree. PAUP uses a number 'seed' to determine the initial ordering: simply by having PAUP run the analysis over and over, choosing a different randomized number seed each time, one can generate many different trees. PAUP can then compare these trees to determine a 'confidence level' for each branching: this is called 'bootstrap analysis'. For example: we ran a bootstrap analysis on the Nun's Priest's data using the neighbour joining method to create 100 different trees, from 100 different number seeds. We learnt from this that all 100 trees support the long-recognized a manuscript grouping, of Dd Ma Cn Ds1 En1; however the grouping Cp La Lc Mg Sl2 Dl, all in Manly and Rickert's c and d groups, is supported by only 57 of the 100 trees.

PAUP offers many different ways of viewing the trees it finds. We have found the following useful:

1. For individual trees made by neighbour joining or parsimony: an unrooted cladogram. This normalizes the distances between each witness, so presenting a clearer tree, though at the cost of making all witnesses appear as if they are separated by the same amount of variation. As it is unrooted, the tree does not appear to grow down or up from a single point: rather it seems to grow outwards by branchings from an approximate centre, which is usually not identified with any one witness
2. For individual trees made by neighbour joining or parsimony: an unrooted phylogram. This retains the distances between each witness, so that it is possible to see where (for example) a particular witness is very

similar to or very different from those close to it. This is particularly useful for assessing how much evidence there is for a specific grouping: well-supported groupings will be separated by a relatively long branch from the other witnesses.

3. For consensus trees derived from bootstrap analysis: a rectangular tree. This shows clearly the number of trees supporting each group, but the groups themselves are less evident due to the very linear presentation.
4. For consensus trees derived from bootstrap analysis: an unrooted tree. This shows the groups rather more clearly, but the number of trees supporting each group is less evident.

1.2 The PAUP analysis of this data

A single NEXUS file, containing the full data for The Nun's Priest's Tale and all three links, was created from the same XML collation output presented in this publication. This collation was made using the parallel segmentation method described in the article 'Rationale and Implementation of the Collation System' included on the Miller's Tale CD-ROM, and available online at http://www.canterburytalesproject.org/pubs/MI-collsystem.pdf. This NEXUS file is NPall.nex, contained in the folder 'Nexus data', contained in the 'NP' folder in the Anastasia distribution folder on your CD-ROM. The ASSUMPTIONS block of this file contains the following:

1. ancstates archetype=0: all; – that is, declaring that '0', typically the reading of the Hengwrt manuscript, is the ancestral state. This could be used to aid tree rooting, but as we are making unrooted trees this has no effect
2. weights scale/basewt=1:all; – thus, all variant places weigh precisely '1'. The effect of this is that at a single place where there are many variants, each variant has correspondingly less weight compared to a place where there are fewer variants. Thus in line 1 of The Nun's Priest's Tale: there are two variants on the third word, widwe/woman and so each weighs 0.5 (1 divided by two); there are four variants on the sixth word (stape/steppid/y stept/y stept []) and so each weighs 0.25 (1 divided by four).
3. various charsets are declared: these permit analysis of different sets of variants (thus, all variants in L30; or only one word variants, etc.)
4. various taxsets are declared: these permit analysis of different groups of witnesses, arranged according to the four different segments of the three links and the Tale. One can therefore readily select the corresponding set of variants and witnesses for any of these segments (thus, to analyze just the Tale, select the charset NP and the taxset NP by the **Include-Exclude Characters** and **Delete-Restore Taxa** items of **Data** menu in PAUP.)

I state above that in this textual tradition a minimum of some 500 'characters' (in the language of evolutionary biology: each character is a single place of variation, typically all the variants on a single word here) is necessary for reliable results. Links 31 and 32 have respectively 118 and 50 characters and so no analysis was carried out on these. There is more data for Link 30, with most of the manuscripts present, and a total of 438 characters, close to the 500 threshold. However, Link 30 falls into two distinct sections: lines 1-4 and 5-34, present in all 49 witnesses, and the twenty lines numbered 4-1 to 4-20 in our system, present in Ad1 Ad3 Bo1 Bw Ch Cn Cx2 Dd Dl Ds1 El En1 En2 En3 Fi Ha2 Ha4 Ht Ii La Lc Ld1 Ln Ma Mg Mm Nl Ph2 Ph3 Pn Ry1 Ry2 Sl1 To1, and so not present in Cp Cx1 Ha3 He Hg Mc Me Ne Ps Pw Py Se Sl2 Tc1 Tc2 Wy. After this division, each of these sections has respectively 276 and 162 characters, again insufficient for reliable analysis on its own.

This left only the set of data for the Tale itself (5222 characters). This set of data was processed four times by PAUP, with each process producing a distinct set of results, as follows:

1. Neighbour joining trees. For this we chose the **Distance** method, selected 'Constrain branch lengths to be nonnegative' and 'Collapse branches of effectively zero length when searching' from the **Distance Settings** option. We then selected **Neighbor Joining/UPGMA**, and chose the **Neighbor joining** option. All these commands are to be found under the PAUP **Analysis** menu. All other settings were kept to the PAUP defaults. The witnesses (taxa) were joined systematically, that is in the order in which they are presented in the NEXUS file. This analysis produced a single tree. This tree is shown as a phylogram in the file 'NP-phyl-NJ.gif' and as a cladogram in the file 'NP-clad-NJ.gif'; the file 'log.txt' gives the PAUP log for this process (all within the 'NJ-NP' folder).
2. A neighbour joining bootstrap analysis. For this we used the same neighbour joining settings described in (1), but this time, PAUP used a randomized number seed to determine the order in which the taxa are joined. By

using a 100 different number seeds PAUP found 100 different neighbour joining trees. The bootstrap analysis then created a consensus tree from all these (that is, retaining the groupings best supported among the 100 trees) and then shows how many of the 100 trees support any given grouping. This tree is shown as a rectangular tree in the file 'NP-rect-NJ-BS.gif' and as an unrooted tree in the file 'NP-unroot-NJ-BS.gif'; the file 'log.txt' gives the PAUP log for this process (all within the 'NJ-bootstrap-NP' folder).

3. A maximum parsimony tree. For this we chose the **Parsimony** method, and chose **Heuristic Search**, selecting 'random' and '100' replications ('reps') from the 'Stepwise-addition options' box. All these commands are to be found under the PAUP **Analysis** menu. All other settings were kept to the PAUP defaults. This analysis produced a single tree. This tree is shown as a phylogram in the file 'NP-phyl-Pars.gif' and as a cladogram in the file 'NP-clad-Pars.gif'; the file 'log.txt' gives the PAUP log for this process (all within the 'Pars-NP' folder).
4. A maximum parsimony bootstrap analysis. For this we used the same parsimony settings described in (3), with the exception that PAUP went through ten replications each time rather than 100. This time, PAUP saved the 100 most parsimonious trees it found, rather than just saving the most parsimonious. The bootstrap analysis then created a consensus tree from all these (that is, retaining the groupings best supported among the 100 trees) and then shows how many of the 100 trees support any given grouping. This tree is shown as a rectangular tree in the file 'NP-rect-Pars-BS.gif' and as an unrooted tree in the file 'NP-unroot-Pars-BS.gif'; the file 'log.txt' gives the PAUP log for this process (all within the 'Pars-bootstrap-NP' folder).

The search results and all the data for these searches are contained in the folder ‘Nexus data’, contained in the ‘NP’ folder in the Anastasia distribution folder on your CD-ROM. This folder contains:

NPall.nex	*file containing the full nexus format information for the Tale and the links*
Driver.xml, L30nex.xml, L31next.xml, L32nex.xml, and NPnex.xml	*files containing all the XML apparatus from which these analyses derive; driver.xml unites them all into a single XML document*
NJ-NP	*folder containing the neighbour joining analysis for The Nun's Priest's Tale*
NJ-bootstrap-NP	*folder containing the bootstrap analysis using neighbour joining for The Nun's Priest's Tale*
Pars-NP	*folder containing the parsimony analysis for The Nun's Priest's Tale*
Pars-bootstrap-NP	*folder containing the bootstrap analysis using parsimony for The Nun's Priest's Tale*
vargroupprofiles.html	*file containing the full variant group profiles for all three links and for the The Nun's Priest's Tale. See below.*

The discussion hereafter is based on the tree created by the parsimony analysis for The Nun's Priest's Tale. Hereafter, we refer to this tree as the NP tree. You may view this tree as an unrooted phylogram and as an unrooted cladogram. You may also like to compare this tree with:

- The tree made by neighbour joining: as an unrooted phylogram and as an unrooted cladogram
- The bootstrap analysis consensus tree made with neighbour joining: in rectangular form and in unrooted form
- The bootstrap analysis consensus tree made with parsimony: in rectangular form and in unrooted form

2. The witness groups in The Nun’s Priest's Tale

The creation of hypothetical trees of relationship as described in the previous section is only the first stage of our analysis. The next stages are as follows:

1. We identify what we call the ‘fundamental witness groups’: groups of witnesses which appear to descend from a single ancestor below the archetype
2. We then identify the variants characteristic of each group. If the fundamental witness groups really exist and there really was a common ancestor to all the witnesses in the group, then we will be able to identify the variants likely to have been introduced by that common ancestor. We use the VBase program (now part of the Anastasia publishing system) to search for these.
3. We then assess the variant groups by closer study of the variants themselves. Do the variants constitute a

genuine genetic group; that is: are they likely to have been introduced at the making of a single copy and then descended into the copies of that copy? Or are they simply the result of chance agreement? Are the variants likely to be authorial, and so either vestiges of Chaucer's original or the results of his revision, or scribal?

4. We then study the distribution of the variants across each witness and across the whole tradition. Is their evidence of contamination, shift of exemplar, or other lateral transmission? What are the affiliations of each witness?

2.1 Identifying the groups

Our initial hypotheses concerning the witness groups are based on the NP tree created by parsimony analysis.

It appears from the NP tree that the witness groups are as follows:

a: The pairs Cn Ma (with Ln) and En1 Ds1 with Dd

b: Cx1 Tc2 Ne. Because of the evidence of correction in Cx2, and hence its descendants Pn Wy, these are excluded from this group for this analysis

c: Cp La Sl2 with Mc Tc1 and the pair Lc Mg

d1: Fi Ra2 To1

d2: Ha2 En2 Mm Ph3 Pw Ry2 Sl1

e: The pair Bo1 Ph2

In addition to these the following manuscripts appear grouped in the NP trees:

o: the trio Hg El Ch with Ad3 and possibly Py Ha3 Gg Ld1 Ha4

However, as the discussion below indicates, there is reason to think that while this last is shown as a group within the tree, the manuscripts within the group may not have a common ancestor below the archetype. To put it another way: their only common ancestor may be the archetype itself. Hence, their grouping in the tree may be an artifact based only on their shared archetypal readings.

Note that one may gain a sense of the support for a particular grouping from the distances between nodes and witnesses on the phylogram. The scale ‘50 changes’ gives an approximation to the number of changes occurring along each branch. For example: the existence of b as a separate grouping, with a common ancestor below the archetype, can be seen from the length of the branch running to the common root of the clusters around Cx1. The length of this branch, including the section up to the rooting point of Hk, corresponds to around 200 variants. As we will see later, variant database analysis suggests that there are some 200 readings characteristic of the b group and thus are likely to have been introduced by the common ancestor of b. By reverse: the shortness of the branches between the common ancestors of the c and d2 witnesses, that is, between the node uniting Cp La Sl2 and that uniting Pw Sl1 Ha2 En2 Ry2 Mm Ph3 suggests that there are rather few variants separating the two groups. Indeed, variant database analysis suggests that there are only 26 readings introduced by the common ancestor of d2. In a variant of this: the incunable edition Cx2 (Caxton’s second edition) is shown as coming from a very short branch from the node it shares with the two later incunable editions, those of Wynkyn de Worde (Wy) and Richard Pynson (Pn). In fact, we know that these later editions were actually copied from Cx2, and the short branch to Cx2 is consistent with this. Similarly, the very short branch between Cx1 and the node linking it with Tc2 suggests that Tc2 may be copied from Cx1.

Close scrutiny of the NP tree suggests the following relations among the fundamental groups:

1. It appears that c and d1 d2 share an ancestor, the cd ancestor
2. It appears that e and cd may share an ancestor. The approximate position of this ancestor is marked with γ on the NP tree.

Previous analyses of sections of the *Tales* published by this project (that is: of The General Prologue, Miller's Tale and Wife of Bath's Prologue) have suggested that there are other 'higher-level' groups uniting smaller individual groups. The analysis of the NP tree did not support the existence of these higher-level groups:

1. ab: analysis of other sections suggested that the a b groups might share an ancestor below the archetype, the ab ancestor. But the tree seems to suggest that the two groups descend separately from the archetype
2. d: analysis The Miller's Tale suggested that the d1 d2 groups might share an ancestor below the archetype, the d ancestor. But the tree seems to suggest that the two groups descend separately from the γ archetype
3. α: analysis of other sections suggested that the ab group might share an ancestor with several other manuscripts (notably Ad3) below the archetype, the α ancestor. But the tree seems to suggest that the not only is there no ab archetype, but that Ad3 and other putative α (Ad1 En3; perhaps Ha4 Gg) do not share an exemplar below the archetype

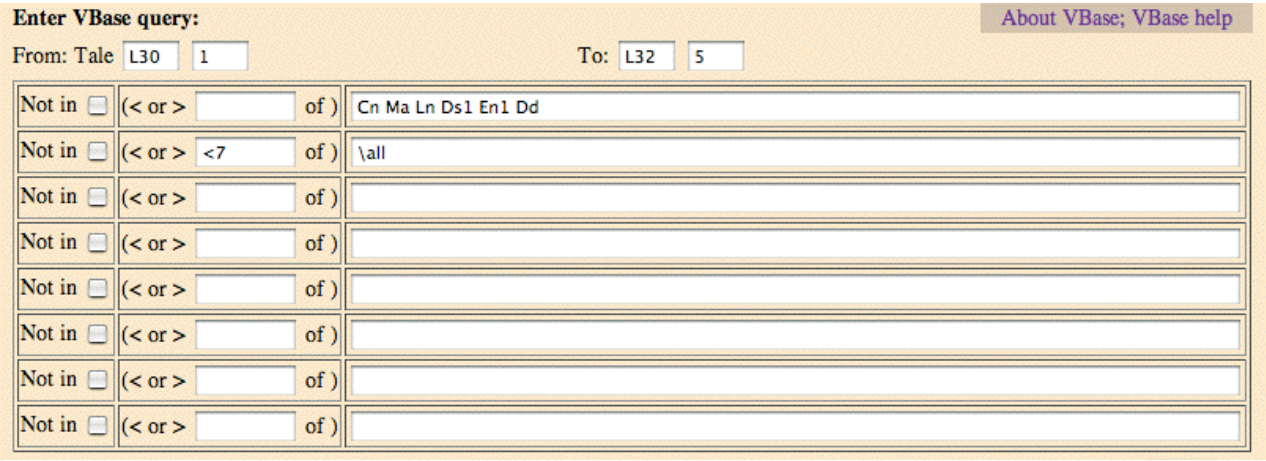
However, we will retain these possible higher-level groups within the analysis, and seek further evidence to confirm or deny their existence.

This discussion suggests that we may be dealing with eleven different witness groups: α ab a b γ cd e c d d1 d2. In addition, we are interested in the variants shared by the o witnesses, and in the alignment of all manuscripts against the single most important division within the tradition, that between El and Hg. Accordingly we should like to know where manuscripts agree with Hg against El ('Hg not El') and with El against Hg ('El not Hg').

2.2 Identifying the variants characteristic of each group

The next step is to identify the variants characteristic of each group. For groups arising below the archetype, these will be the variants most likely to have been introduced by their common ancestor below the archetype (technically, their 'hyparchetype'). We use the variant database facility, or 'VBase', built into the publication interface to identify these.

Identifying these groups of characteristic variants is not straightforward. Consider the case of the a witnesses. This group appears to comprise Dd and the pairs Cn Ma (here with Ln) and Ds1 En1. As explained at length on the General Prologue CD-ROM, one cannot find the variants characteristic of a by simply asking: find all the variants in these six witnesses and nowhere else. Indeed, you can test this for yourself by running the following query in VBase:



This search produces just seven hits: the replacement of 'that' by 'which thathe' in the phrase 'Or an Agew that may be youre bane' in NP 140; the omission of 'bothe' in NP 171, and so on. But seven hits would not be enough upon which to build any kind of argument.

Such searches presume that the variants characteristic of any one group appear in all the witnesses of the group, and in no other witnesses at all. The reason searches built on this model do not work is because this makes three assumptions:

1. A variant introduced by a common ancestor will be retained in each and every descendant of that common ancestor. That is: every reading introduced in the A ancestor will be present in every descendant of that

ancestor

2. A variant introduced by one copyist at one point in the tradition will not be introduced independently by any other copyist at that point
3. Once a variant is introduced into the witnesses of one group, no scribe copying a witness of a second group would deliberately import the variant from the first group into the copy

All three assumptions are manifestly false:

1. For the first: if a variant can be introduced in one copy, it can be removed by any one of the descendants of that copy, either by the descendant introducing a further variant, or replacing the variant by the original reading - or, the descendant might lack the whole passage. The larger the number of descendants the more likely this is to happen
2. For the second: as Talbot Donaldson puts it, there is little originality in sin. If one scribe makes a mistake, it is very likely that some other scribe somewhere will find just the same way to make a mistake (1970, 108). The larger the number of witnesses in the whole tradition the more likely it is that any two scribes (or more) will make the same mistake, quite independently. Manly and Rickert called this 'agreement by coincidence' (ACCO; see II 20-21); evolutionary biologists call it convergent evolution.
3. For the third: the literature is rich in copyists behaving as eclectic editors. Sometimes they will deliberately compare two versions, introducing the readings of the second into the copy of the first (as Caxton did in preparing his second edition, Bordalejo 2002). Sometimes they will remember a striking variant and introduce that. Both these are instances of contamination; evolutionary biologists call this hybridization or lateral gene transfer. Again, the larger the number of witnesses the more likely that such transfers will happen.

The effect of these considerations is that we have to abandon the absolutist approach of accepting that only variants found in all the witnesses of a particular group and only in those witnesses can be evidence that they share a common ancestor.[1] In its place, we have to adopt a probabilistic approach, as follows. If a variant is likely to have been introduced by the common ancestor of a hypothetical grouping of witnesses, then the following is likely to be true of that variant:

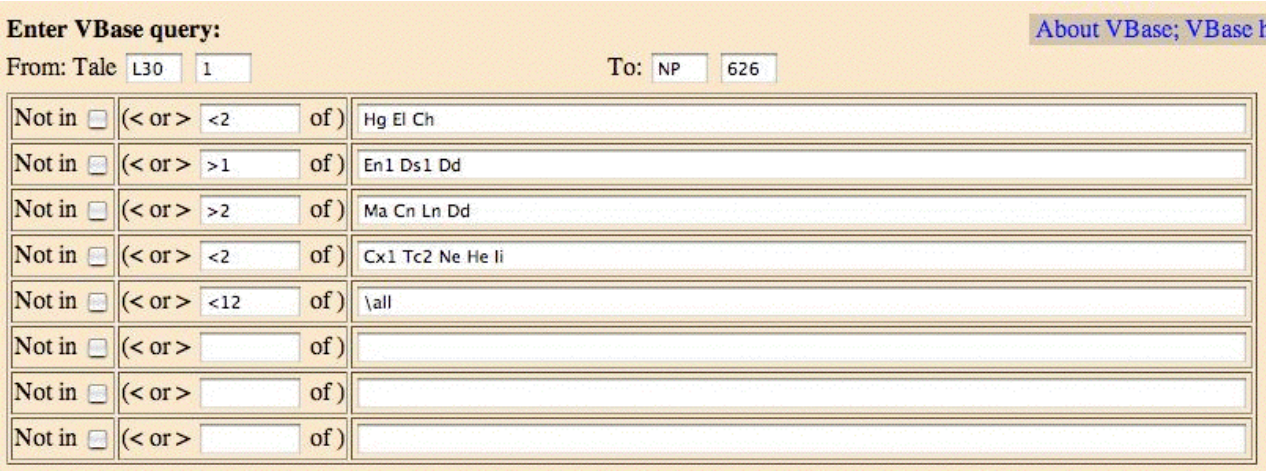
1. The variant was probably not present in the ancestor of the whole tradition.
2. The variant is present in a significant number of witnesses within the hypothetical grouping
3. The variant is not present in a significant number of witnesses outside the hypothetical grouping, and especially not in witnesses in groupings closely related to the group being analysed.

For each of these:

1. If the first is not true: then the variant will not have been introduced by an ancestor below the archetype, and its presence in any witnesses at all is only evidence that they are descended from the common ancestor of the whole tradition - precisely, evidence of nothing.
2. If the second is not true, then obviously we are wasting our time.
3. If the third is not true, then either the variant was introduced at a level some way above the common ancestor of this group (and so is evidence of a wider grouping than just this), or is so widespread that it is either ancestral to the whole tradition or so easily introduced as to be commonplace, and, again, evidence of nothing.

The identification of varants which satisfy these conditions requires a rather complex search. One is looking for variants which fit a pattern such as: find all variants in more than one of witnesses X Y Z, and in more than one of A B C, and in less than two of D E F. The VBase tool is designed for exactly such searches.

Consider again the case of the a witnesses, hypothetically Dd and the pairs Cn Ma (in NP with Ln) and Ds1 En1. The VBase search for this is as follows:



The first line aims to rule out the variants which do not satisfy the first criterion, that is, those variants present in the ancestor of the whole tradition. The three manuscripts El Hg Ch appear to be the three nearest to the lost **O**, the archetype behind all surviving manuscripts, and may represent three independent lines of descent from it. If a variant is present in any two of these, it is highly likely that it was present in **O**, the ancestor of the whole tradition. (See also Bordalejo's discussion of the a group in the "Witness Groups" article on this CD-ROM.)

The second and third lines aim to identify the variants actually characteristic of this group, and so satisfy the second criterion. The group consists of the single manuscript Dd, the trio Ma Cn Ln and the pair En1 Ds1. It would not be enough for a variant to be present in Dd alone; nor would it be enough for it to be present in just one of the groups Cn Ma Ln or En1 Ds1. It should be present either in the manuscripts of one group, or in one manuscript of one group and Dd, and also present either in the manuscripts of the other group, or in one manuscript of the other group and Dd.

The fourth and fifth lines seek to eliminate the variants which do not satisfy the third criterion: that is, the variants most likely to have been introduced by the immediate ancestor of the group and not to be variants particularly likely to arise at any point of copying. The fourth line specifies that it should be in no more than one of the b witnesses (disregarding for this purpose the incunables Cx2 Wy Pn, where contamination in Cx2 might muddy matters). The phylogenetic analysis for other tales (though not for NP) suggests that a and b might share an ancestor. Therefore, if the reading were present in more than one b witness it is likely that it was introduced in the ab ancestor (or higher), not by the a ancestor. However, it might be in one b witness just by coincidence and still have been introduced by the a ancestor. The fifth line specifies that the variant should not be present in twelve or more witnesses. This is a 'safety valve': the likelihood is that if the variant were present in so many witnesses, then either it is ancestral to the whole tradition or is specially likely to have been introduced at random into quite unrelated witnesses. In either case, the variant will have no power to distinguish the a group and so will have no value for analysis.

Running this query in VBase results in a total of 62 hits for NP.

3. Assessing the variant groups

Fifteen separate groups of variants were created using the methods detailed above. These were:

- The eleven groups a b ab α c d1 d2 d cd e γ corresponding to groups apparently revealed by the NP trees, or presumed from analysis of other sections of the *Tales*
- The two groups 'Hg not El' and 'El not Hg', giving for each manuscript a statement of how many agreements it has with respectively Hg against El and with El against Hg
- The group 'Cx2', comprising the variants found not in Cx2 but not in Cx1, and so likely to include readings imported by Caxton into his second edition from the now-lost better manuscript he used while preparing this edition
- The o witnesses: though not a group, in the sense that they may not share a hyparchetype (a shared ancestor

below the ancestor of the whole tradition), yet the variants characteristically found only in these may be of considerable interest

You may view the variants in each group by selecting the link from the 'Search VBase' page (accessible by selecting from the pop-up menu on the right in the navigation toolbar).

We then assessed each group of variants by examining them, for two factors:

1. Is the character of the variants such as to suggest they might be authorial, and therefore either ancestral to the whole tradition or potentially derived from a revision by Chaucer, surviving in the manuscripts of this group?
2. Is the character and distribution of the variants such as to suggest that they might be the result of coincident variation, occurring at many times within the tradition rather than of a single act of copying introducing a discrete set of variants?

For all but two of the thirteen groups (setting aside the 'Hg not El' and 'El not Hg' groups) the answer to both questions is clearly negative. In the a group, for example, observe the following variants:

NP 14 Of poynaunt sawce **hir neded** neuer a deel (24 witnesses)
a: Of poynaunt sawce **knewe she** neuer a deel (6 mss including 5 of the a group)

This is an obvious case of substitution of a more obvious, easier reading for a more difficult one, and one more likely to be authorial. The presence of the readings in 5 of the a group suggests this is not an accidental agreement in a - as appears to be the other instances of this reading, in Nl.

NP 94 To han **housbondes** hardy wise and fre (45 witnesses)
To han **an housbonde** hardy wise and fre (7 witnesses)

The substitution is routine, and appears in just one witness beside the six Dd Cn Ln Ma En1 Ds1.

NP 103 Sweuenes **engendren** of replexions (31 witnesses)
Sweuenes **been engendred** of replexions (10 witnesses)n

Again, a substitution of an easier reading for a harder reading and, again, appears in all six a witnesses and just four outside the group.

Analysis of the variants in eight of the groups in NP suggest that the groups arose below the ancestor, through non-authorial copying from a single hypearchetype which introduced the readings characteristic of each group. The eight groups are a b γ cd e c d1 d2.

Among these eight, the pattern seen elsewhere in the *Tales* of several of these groups being related to one another holds good. Here, we find the same pattern as we found in The Miller's Tale, that the five groups cd e c d1 d2 (incuding among them more than half the manuscripts of the entire tradition) all appear descended from a single copy below the ancestor. We have called this single ancestor γ. The confirmation that all these manuscripts for both these two tales, which are widely separated in all the manuscripts, share the same textual history is striking confirmation of the coherence of the history of the tradition.

Analysis of other parts of the *Tales* has also suggested that the a and b groups share an ancestor below the archetype. The NP analysis agrees with analysis of each sections, suggesting that the two groups a b persist in NP, each apparently descended from a single ancestor below the archetype and so each marked by a distinctive set of variants, apparently introduced by the ancestor of each. However, for the NP the evidence of the phylogenetic analysis is weaker for a shared ancestor of the the two groups below the archetype. Indeed, analysis of other sections (particularly, The General Prologue) has suggested that there may be an ancestor – the so-called α ancestor – above ab, shared with other manuscripts below the archetype (notably, Ad3 Ha4 Gg). There is weaker evidence of this α ancestor in The Nun's Priest's Tale also. This does not mean that the there were no ab and α ancestors for NP. It could be simply that both these introduced many fewer changes into their copies of NP than they did for GP amd MI, and so left little trace in the record.

4. Issues in the textual tradition of the Nun's Priest's Tale

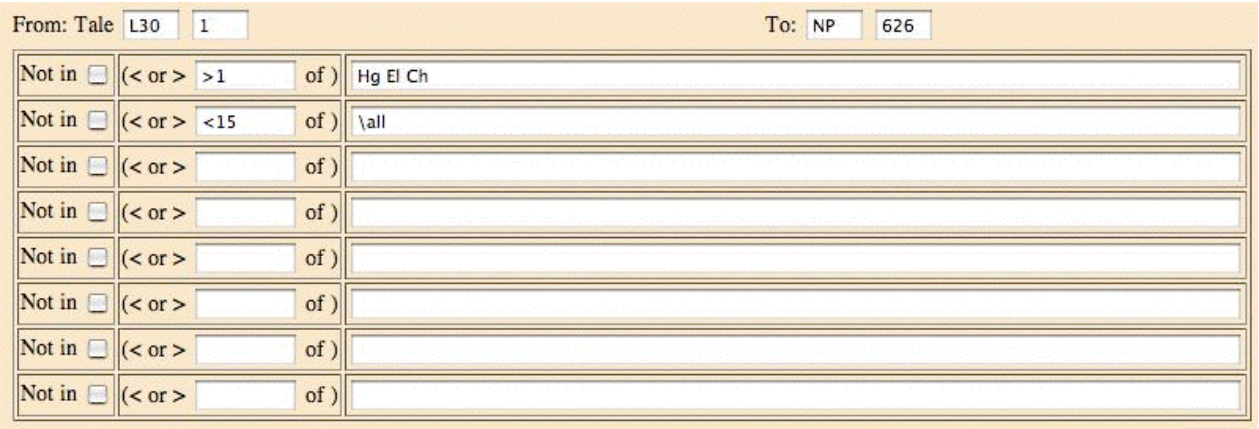
4.1 The o witnesses

VBase search for variants characteristic of these witnesses

Analyses of other parts of the *Tales* have identified sets of witnesses which do not appear to belong to any of the large groups (a b γ) identified in the last section, and which do not appear to share an ancestor below the archetype with any other witnesses. Effectively, these may all descend independently from the archetype. The test for o is very simple. We presume that from the closeness of the three witnesses Hg El Ch, as shown by phylogenetic analysis, and from the near-uniform excellence of the readings in those witnesses, that any reading in at least two of these three is likely to have been in the archetype of the whole tradition. However, if these readings are found in the great majority of the witnesses, they will be of little interest: simply, they will be readings which have descended everywhere. But readings present in at least two of these three (and so likely to have been in the archetype) but which are not found throughout the whole tradition, will be of considerable interest. This situation might arise for one of two reasons:

- There is something about the reading which makes it particularly susceptible to alteration or loss, as lectio difficilior
- Or: the reading of two of Hg El Ch might be an error shared by these, with the other witnesses actually preserving the archetypal reading

In either case, the reading and its analysis will be illuminating. Thus, we carried out the following search using VBase:



This yielded 67 readings. Most of these appear of authorial character, so belonging to the first category given above. The very first of these, in L30 6, is typical:

> L30 6 Wher fore sire monk **o** Daun Piers by youre name (11 witnesses: o 4 witnesses [Hg Ch Py Ha3]; a 4 witnesses with Dd Me2; Wy)
>
> > Wher fore sire monk **or** Daun Piers by youre name (10 witnesses, with other variants)
> > Wher fore sire monk [ ] Daun Piers by youre name (27 witnesses, including El Ad3 Ld1 Ha4, with other variants)

Pearsall is the first editor to choose the Hg reading, 'O'. He justifies this entirely in terms of poetic and dramatic effect: 'a self-consciously extravagant kind of ingratiation by the host...[t]he whole speech is highly dramatic.' (Pearsall 1984, 132-33). It appears from Pearsall's collation that Hengwrt is the only manuscript to have this reading. Indeed, of the ten manuscripts collated for the *Variorum* by Pearsall only one other (Dd) has this reading, though Pearsall does not record this. However, our full collation shows the reading as present in eleven witnesses, and the reading can also be supported by our stemmatic analysis. Firstly, it is present in three other o witnesses,

usually independent of Hg: Ch Ha3 Py. Secondly, it was clearly present in the ancestor shared by the four a group manuscripts and both Dd and Me. Thirdly, its presence in Wy, which appears to use for the link a now-lost manuscript (see below), may be evidence of its presence in yet another line of descent. Thus, these eleven witnesses may actually represent no less than six lines of descent (one for each of Hg Ch Ha3 Py; a and Wy). Against this, 17 of the 27 witnesses which omit 'or' represent just one line of descent, from γ. Thus, despite the disparity in numbers,the rather stronger support within the o witnesses for the Hengwrt reading weighs in favour of Pearsall's choice.

Similarly, though less dramatically: o preserves a harder, more unusual reading, widely replaced elsewhere by a more commonplace reading in L30 14 (as that->as), NP 22 (loss of double negative), NP 29 (heet/hight->that hight: note that the trio Ch El Hg have the spelling 'heet', unique to these three: one of the straws of evidence suggesting that Ch might be copied from an exemplar written by the Hg/El scribe), NP 35 (krew->knew), NP 57 (was it->it was: Pearsall 1984, 154 describes the line as 'characteristically Chaucerian'), NP 81 (han had->wolde han had), NP 108 (colera->colour/coleryk), NP 114 (ful many->many/wel many), NP 115 (blake->grete; or->and), NP 160 (as->as of), NP 162 (make->to make/nouht to make), NP 170 (o->a), NP 216 (wente as it were->as he wente/variants) NP 217 (that same->the same), NP 222 (heere he lyth gapyng->he lyth gapyng: Pearsall 1984, 184 comments on the metrical awkwardness, but acceptability, of the o reading here), NP 237 (this my->this is my; see Pearsall's note), NP 238 (that town->the town), NP 247 (see->the see), NP 256 (herkneth->herkne: Pearsall 1984, 188 again justifies the metrical awkwardness on poetic grounds), NP 264 and 265 (to lette->for to lette; byde->abyde; again the o readings are metrically awkward, Pearsall 1984, 189), NP 272 (or->and), NP 283 (it->ther/hym/hem), NP 293 (lite->litel, replacing an unusual form), NP 297 (was->nas) hadde->various replacements), NP 346 (al his->his/mannes), NP 356 (he was->and was/and/[], NP 361 (foot->feet), NP 362 (He->And/Ay), NP 363 (rennen thanne->thanne ran), NP 395 (of sley->of sleight and/and of), NP 398 (forncast->aforncast), NP 407 (newe Genyloun->and newe Genyloun), NP 411, NP 428 (er that I->er I), NP 434 (that->the), NP 440 (to whom->whom; it->I), NP 453 (bifel that as->bifel as), 459 (flee->to flee), NP 464 (wher->what; see Sisam, 51), NP 468 (for tespie->to espie; Pearsall 1984, 228 is incorrect in asserting that all the Variorum collated manuscripts read 'to aspie'), NP 470 (oonly for to herkne->oonly to herkne/variants), NP 476 (gentilesse->gentilnesse), NP 477 (to myn gret->various), NP 486 (he wolde->he moste/he dide/wolde he), NP 497 (nys->is), NP 517 (ne was ther->was ther), NP 525 (world->the world), NP 555 (The->This), NP 557 (at->at the), NP 568 (hem thoughte hir herte->various), NP 587 (if that->that), NP 600 (to yow quod he ydon->various), NP 603 (wikke->wikked), NP 619 (and->and of a/and a/various), NP 621 (seith that->seith), NP 622 (it is ywrite->it is writen). Taken together, the concentration of these readings in relatively few witnesses close to each other at the centre of the NP trees is very striking, as is the scarcity of instances of coincidental agreement in these readings with witnesses outside these few. It seems that not only did scribes frequently fail to comprehend these readings: once a reading had been replaced by a more commonplace variant few scribes had the wit to restore it. Thus, over and over again, we find a reading preserved by almost every one of the o witnesses, yet lost almost everywhere else: examples are NP 108, NP 395, NP 397, NP 428 (er that), NP 440, NP 459, NP 468, NP 470, NP 476, NP 517, NP 525, NP 600, NP 602, NP 603, NP 621, NP 622 (with c)

All these readings appear to have been present in the archetype, and to have been of a nature that made them particularly liable to alteration by scribal copyists. Pearsall, for example, accepts every one of these readings into his edition. However, in a few cases, it has been suggested that the reading of o (that is, of at least two of Hg El Ch) is an error, while the archetypal reading is found in other witnesses:

- NP 35: all editors have accepted 'he knew' in NP 35, over 'he krew', found in all three of Hg El Ch, in Py and Ha3, in all b witnesses, and in a single γ witness, Pw. While the argument from sense and from the analogues for 'knew' remains strong, 'krew' has strong stemmatic support, with the numerical majority for 'knew' coming from the a and γ witnesses alone.
- NP 411: El and Ch here read 'that yerd', so reading the line as 'That thow in to **that yerd** flaugh fro the bemys'. The Hg reading 'the yerd' (so 34 witnesses, including all a and γ, though only HA4 and Py within o) removes the awkward repetition of 'that', and – as Pearsall 1984, 217 points out – appears to undermine the emphasis on time ('that morwe' in the previous line) in the couplet.
- NP 428: here, Ch and Hg read 'I was wroght', so reading the line 'Though god forwoot it er that I was wroght',

in place of El's 'it was wroght'. Many editors prefer the El reading, but Pearsall 1984, 220 argues strongly on poetic grounds for the Ch and Hg reading 'I'. This reading is also well supported within o, with Ad3 Gg Ha3 Py all reading 'I'.

- NP 551: Hg reads 'Of Rome cryden the senatours wyues', and this is the reading of every manuscript outside o and the a group (39 witnesses). But within o Hg is supported only by Py Ad3 Ha4: all other o witnesses for NP (El Ch Dd Ha3 Gg Ad1 En3) and the whole of a omit 'the'. On this occasion, Pearsall (who usually supports the Hg reading, which he does here retain in his text) acknowledges both the metrical superiority of the El reading and the likelihood that scribes, failing to recognize this, might independently insert 'the'.
- NP 566: a classic crux. Hg reads 'So fered for berkynge of the dogges'. The syntax of 'So fered' is awkward, but the strong support for it within o (all three of the key Hg El Ch, Dd, a, with Ad1 En3 Py also all reading 'So fered') makes it highly likely that this was the reading of the archetype. The scribes clearly found the phrase as difficult as modern editors, with our collation identifying thirteen variant phrases here.
- NP 584: Hg reads 'The hope and pryde eek of hire enemy'. Even Pearsall 1984, 251 admits embarassment at the Hg inclusion of 'eek' here, apparently as a metrical 'improvement', and expresses a yearning for the El reading, which omits 'eek'. But El is quite alone here, with the spread of variants across the tradition indicating both the likelihood that the archetype included 'eek', and the difficulty the scribes had with this reading.
- NP 602: o is near unanimous in reading 'Whan I yow hente and broghte into this yerd'. Yet, this appears to make no sense: the fox has taken Chauntecleer out of the yard, not into it, and so almost all scribes outside o and almost all modern editors read 'out of'. But Pearsall 1984, 253 defends the o reading (which also seems to have stood in the a ancestor) as instance of 'Chaucerian subtlety': 'The desperation is the fox's, not the text's'.

The striking quality of these variants, together with their vulnerability to scribal simplification and the rarity of their restoration, suggests that Chaucer himself was responsible for these readings. Their co-occurrence in just a few manuscripts near the centre of these trees (Hg Ch El, of course, but also in significant numbers in Ad1 Ad3 Dd En1 Gg Ha3 Py) argues in turn the closeness of these witnesses to the archetype of the whole tradition. Outside these o witnesses, we should note the high number of these o variants in manuscripts of the a group (Cn Ma En1 Ds1 with Ln). It appears that the a hyparchetype was itself an o manuscript, and very close to Dd.

It follows that, as the o variants are archetypal, their occurrence in these or any other witnesses does not suggest they have a common ancestor below the archetype. These witnesses are of exceptional interest, because of the likelihood that at many points they preserve the most original text. But they do not form a group in the same way as a b and the others. One might speak of o witnesses as shorthand for these witnesses, and one might speak of o variants denoting the archetypal variants clustered in these witnesses. But one should not speak of an o group of witnesses; strictly, there is no such thing.

4.2 The position of the c group within the tradition

VBase search for variants characteristic of these witnesses

It has been long recognized that a large group of manuscripts, usually represented by Cp and La and labelled c by Manly and Rickert, appear descended from a distinct hyparchetype. Because of the early date of Cp, particularly, one might expect to see this ancestor as close to the archetype. But the phylogenetic analysis shows that Cp is quite distant from the archetype, which is likely to have stood close to the Ch El Hg cluster. Further, several groupings usually seen as derived from c, the groups we label d1 and d2, in this analysis appear to descend from an ancestor above Cp and the c hypparchetype.

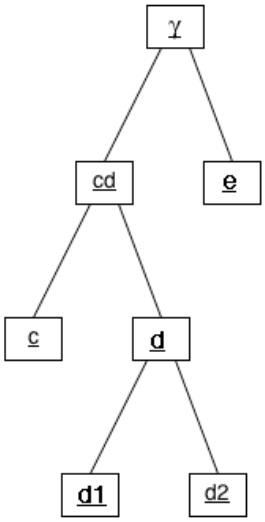
The following are relevant:

- From our analysis of both The Miller's Tale and The Nun's Priest's Tale, it appears that there was at least one copy between the c hyparchetype and the archetype. In both analyses, we call this γ, and we argue that it was the ancestor of the c and e groups. For both The Miller's Tale and The Nun's Priest's Tale, the evidence that the c and e groups share variants below the archetype is sound: VBase identifies some 62 such variants in NP (γ:), 26 in MI. Thus, at least one layer of copying, represented by γ lies between c and the archetype.

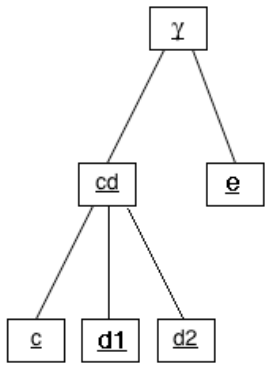
- The relationship between d and c is uncertain. Elsewhere (for example in The Miller's Tale and in The General Prologue) it appeared that d might be descended from a c exemplar. But the disposition of the witnesses in the phylogram suggests that d is actually descended from an archetype above, or at least parallel with, c. There is a possibility that c and d may share an ancestor, apparently below γ: thus the 32 readings identified by VBase shared by c and d but not present in the branch of γ represented by e (cd:

  VBase search for variants characteristic of these witnesses

). This would place yet another layer of copying, represented by this cd ancestor, between c and the archetype.
- It should be noted that d1 and d2 appear to separate in NP. While it is clear that both d1 and d2 are descended from the cd ancestor (with the manuscripts of d1 and d2 all having between 20 and 29 of the 31 cd variants), the evidence for d1 and d2 sharing an ancestor below cd in NP is weak: VBase finds only 16 variants likely to have been introduced by a d ancestor below cd. The same search in MI identified 33 variants, indicating that in MI there does seem to have been a d ancestor below cd. The difference between the two may be represented as follows:



Relationships within γ in MI, showing d1 and d2 sharing an ancestor below cd



Relationships within γ in NP, showing d1 and d2 descending direct from cd

In practice, whether d1 and d2 share or do not share an ancestor below cd makes little difference to our analysis. In any case, one could explain the difference between MI and NP here just by the d ancestor being a more faithful copy of cd in NP than it is in MI. Thus, it exists in both: just that in NP, it introduces rather fewer variants (under twenty only) than it does in MI.

4.3 The links and 'added' lines within the links and the Tale, and the question of Chaucerian revision

There are more than usual difficulties associated with the links and with so-called 'added' lines within them and the Tale in NP. Thus:

- While the headlink (our L30) is present in almost all manuscripts (50 of the 55 which have the Tale), within L30 all but 16 manuscripts include a passage of 20 lines, not present in Hg. This is L30 4-1 to 4-20 in our numbering: B2 3961 to 3980 in the traditional numbering, Riverside VII 2771-2790. Beside Hg, the witnesses which do not have this twenty lines are distributed right across the tradition: in o Hg Py Ha3 Me; all b except Cx2 (the lines were apparently in the Cx2 exemplar used by Caxton) Pn Ii (which has the lines inserted after L30 6); in d2 Pw alone; in c Cp Sl2 Mc Tc1; and the unaffiliated (or contaminated) Ps Se
- Two forms of an epilogue to the Tale are found in the manuscripts. L31 is the 'short' form of the epilogue, and is found in eleven witnesses: the five manuscripts affiliated to the a group (Cn Dd Ds1 En1 Ma); the pairs Cx2/Wy Ad1/En3, and two manuscripts usually independent of these, Ch and Ry1. L32 is effectively an extension of L31, being found only in the a pair Cn Ma and in the o pair Ad1 En3. In these four, L32 follows on immediately on L31, so giving a 'long' form of the epilogue.
- Two lines describing the attachment between Chauntecleer and Pertelote in rather sexual terms are found after NP 50 in ten witnesses: 'He fethered hire an hundred tymes a day/And she hym pleseth alle that euer she may'. The ten witnesses are all five a witnesses (probably also Dd, see Manly-Rickert 2: 423), three incunables (Cx2 Pn Wy) and the unaffiliated Nl Ry1. The occurrence of these lines in Cx2 but not Cx1 is particularly significant: it suggests that the lines were in the manuscript Caxton used in preparing the second edition, which recent analysis suggests was a manuscript of exceptional authority (Bordalejo 2002). These two lines differ from other 'added' lines in their literary character and in their distribution, here appearing in witnesses which other evidence suggests include readings likely to have been present in the archetype.[2]

The occurrence of these lines and passages in some witnesses, alongside others which do not have them, has led scholars to construct narratives of revision by Chaucer. Hammond 1908, 241-2 suggested that Chaucer first wrote the 'short' form of NP 30, later revising it by folding in the additional 20 lines and changing the lines around the added passage to accommodate it. Other scholars have sought to elaborate this by linking this change to the movement of the 'Modern Instances' passage in The Monk's Tale, or to a hypothesis of wider revision within the Tales: Pearsall 1984, 86 summarizes these efficiently.

I do not propose to add to these hypotheses. However, discussion of these, or of any hypotheses of revision of the *Tales* by Chaucer, if it is to be based on evidence, should consider how authorial revision might show itself within the textual tradition of the *Tales*. We can suppose that revision by Chaucer of the *Tales* might have taken one of two forms, and each form would have left distinct traces within the textual tradition:

- Chaucer might have carried out a systematic revision of part or all of the *Tales*, so creating a distinct authorial rescension of the text. This rescension might then have been 'published' separately, alongside a prior, unrevised, form. If this were the case, we would expect evidence in the tradition that a particular witness grouping consistently contained distinct readings which are plausibly authorial.
- Alternatively, the revision could have taken the form of Chaucer changing readings in his own copy, in which case the original and revised readings might be distributed randomly across the tradition, as the scribes chose to incorporate or reject this or that revision. In that case, we might expect to find competing plausibly authorial readings at various points in the tradition.

One can also distinguish two kinds of revision:

- Firstly, Chaucer might have revised, more-or-less methodically, at the word-by-word level. He might have done this either as part of a systematic revision, so distributed in identifiable witness groupings. We have seen no such grouping. Apart from o, all witness groupings are characterized by readings which are plausibly scribal rather than authorial. Within o – for example, just in the pair Hg and El – the patterns of variation so shift as to make it impossible to map the variation into distinct groupings, and then to categorize the distinct groupings as authorial. Or, he might have revised by changing readings in his copy text, which would then be distributed randomly across the tradition, as competing plausibly authorial readings. But so far, at the level of word or phrase variant, we have found remarkably few such competing readings – arguably, none at all. Even

in the o witnesses, exactly the witnesses where one would expect to find such competing readings, we find few or none which cannot be explained by the ordinary processes of scribal copying. One of the few exceptions to this appears to be the variation waspes/whelpe in NP 112: see the stemmatic commentary note ad loc.

- Alternatively, Chaucer might have revised at the 'passage' level: that is by adding, deleting, or repurposing larger or smaller units of text, even whole tales. While one struggles to find a single reading at the word-level which might be convincingly assigned to Chaucerian revision, there are many cases where whole passages – indeed, the whole of the Canon's Yeoman's Tale – are present or not present in different witnesses, or are moved or repurposed within the *Tales* framework. Instances are the so-called 'added passages' in the Wife of Bath's Prologue; the 'modern instances' and the Adam stanza in the Monk's Tale; the whole of the Canon's Yeoman's Tale; differing links around the Clerk/Merchant/Franklin complex – and, particularly, the two forms of L30 and the presence of absence of L31 and L32.

The quality of the text of many of these passages is commonly agreed to be consistent with Chaucer having written them, and their presence in many of the o witnesses is evidence that they were present in very early forms of the text. Prima facie, this suggests that Chaucer did revise at the 'passage' level, developing the *Tales* by adding, removing, or moving line-units at particular places. Once more, he could have done this as a systematic revision, creating a distinct recension of the *Tales* which was separately 'published'. If so, one would expect to be able to identify particular manuscript groupings each embodying different stages of the revision. But no such groupings appear to exist. For example: it is plausibly argued that the 'short' form of L30 (found in Hg and Cp) represents an early version by Chaucer, the 'long' form (found in a) represents a later version. But just a few lines later, after NP 50, a has the couplet 'He fethered hire an hundred tymes a day/And she hym pleseth alle that euer she may', which most scholars argue is an early draft by Chaucer, later rejected by him. Hg does not have this couplet. But how can Hg have both an 'earlier' form of the text than a in L30, then a 'later' form of the text just a few lines later, at NP 50?

Any attempt to explain the distribution of these passages across the witnesses in terms of systematic revision, followed by separate publication of distinct versions, is quickly embarrassed by cases such as the above, where any one 'version' seems inconsistent within itself, containing passages apparently from different layers of revision. However, this distribution is perfectly consistent with the hypothesis that Chaucer revised at the passage level within his working text: that is, by adding passages (in the margin, or on separate slips of paper), or marking passages for deletion.[3] Different scribes might then choose to treat these passages variously, including or excluding according to taste and policy. Broadly, it seems that the policy of the scribe who created the a version was to include such passages, while generally the scribe who created the γ version tends to exclude them. Within o, the witnesses also divide: Hg tends to exclude, El to include. But many factors might disturb any such pattern. The passages might be marked by Chaucer in such a way as to make it more or less likely that scribes might include the passages: this might explain why few manuscripts include the couplet after NP 50, or L31 and L32. Or, indeed Chaucer might have added (or deleted) material from his working copy after one or more copies, or partial copies, had been made. This may well be the case for the two forms of L30: Chaucer might have revised this, adding the twenty-line passage, after Hg was copied, and after the γ ancestor was copied. This would explain the absence of these twenty lines from Hg and the other o witnesses which do not have the lines (Ha3 Me Py), b, the unaffiliated Ps Se, and witnesses usually seen as high in γ (Cp Pw Sl2, here with Mc Tc1).

The lack of these lines in these γ witnesses is another of several pieces of evidence that γ drew on an early state of Chaucer's developing copy: thus the absence of the closing lines of The Merchant's Tale from Cp, corresponding to a change of ink in Hg; the 'chapter numbering' in Cp, suggesting an early attempt at organizing the *Tales*. The puzzle here is not the absence of the lines from Cp and other copies: rather, it is their presence in most of the witnesses descending from γ. One must presume that these lines were added into the γ exemplar after the first copies descended from it were made. This is not so difficult a hypothesis as it might seem. The scribe who wrote Cp, probably the nearest descendant of γ we have, also (almost certainly, later) wrote Ha4, a classic o witness, which included the lines. This scribe would have known the lines from Ha4, if from nowhere else, and could have included them in the γ ancestor, or in another close copy of γ, from which in turn all the later γ witnesses descend. However, further research on the many manuscripts within the γ group (which have been little studied to now)

will modify this tentative suggestion.[4]

4.4 The Hengwrt and Ellesmere copies

For the best part of a century now, the relationship between these two copies and their relative priority in terms of the choices they offer editors and readers have been the most debated issues in the textual scholarship of the *Tales*. To summarize: most editors since Skeat have based their text on El: thus, particularly, the Riverside editions of Robinson and Benson. However, it has long been agreed that Hg is an earlier copy than El, and many scholars have pointed out that at many points Hg has readings preferable to those in El. Manly and Rickert argued that El is in fact the result of a careful revision by someone other than Chaucer, resulting in a general smoothing of metre and language. More recently, Norman Blake has been seen to espouse what has been labelled a 'hard Hengwrtism': if it is not in Hg, it is not Chaucer's.[5].

The Nun's Priest's Tale offers a rich environment for this discussion, for two reasons. Firstly, the two forms of L30 (the Nun's Priest's Headlink or Prologue) – Hg with the 'short' form, El with the 'long' form – present the differences between the manuscripts in particularly challenging form. Is Chaucer responsible for both versions? How did the distribution of the two versions in the tradition arise? Which version should we read – or should we read both? Secondly, Pearsall's *Variorum* edition of the tale (Pearsall 1984) takes a particularly uncompromising approach to the question. He declares:

Any question about the textual authority of other manuscripts in relation to Hengwrt can be answered simply by the Hengwrt editor: they have none. (Pearsall 1984, 4)

So strong an assertion invites contention. Indeed, Pearsall immediately qualifies it: 'other manuscripts' might have 'a subsidiary value' in certain circumstances where Hg is defective or demonstratively in error: however, he seeks to limit this to just 'two or three instances'. But the long form of the link presents a problem, as Pearsall acknowledges that this must represent Chaucerian revision. Accordingly he includes this passage in his edition, discreetly bracketed. In itself, this represents a significant modification of Pearsall's assertion. However, he is elsewhere at pains to stress the inferiority of El to Hg (even as he acknowledges its superiority to every other witness): thus page 98 lists some thirty variants where El is variously in error, inferior, or pedantically revising.

Our analysis confirms Pearsall's argument, with some qualification. I identified sixty-three points, all in L30 and NP, where the readings of Hg and El are directly comparable. These are all places of word or phrase variation: we did not include, for this analysis, whole lines or passages present or not present in either manuscript. Such whole lines or passages may not have been present in both exemplars and so one can make no statement about the copying of either manuscript on the basis of their presence or absence.

These sixty-three places of variation are all discussed in the Stemmatic Commentary in this publication. At each of these places in the Stemmatic Commentary, we attempted to determine (as previously, for The General Prologue and The Miller's Tale) the reading most likely to have stood in the archetype and hence, for all sixty-three, which manuscript preserves the archetypal reading, and which miscopies. I summarize here the conclusions presented in the Stemmatic Commentary about these sixty-three:

- Points where it is not possible to determine which of Hg or El preserves the archetypal reading: 7 (NP 107 [comth/cometh] 160 227 544 551 588 591)
- Points where it appears that El preserves the archetypal reading, and Hg is in error: 5 (NP 36 223 444 547 595)
- Points where it appears that Hg preserves the archetypal reading, and El is in error: 51 (L30 1 6; NP 34 57 58 62 69 74 107 [the grete/grete] 111 122 [2x] 157 158 171 190 216 254 286 299 329 [2x] 357 362 364 394 411 424 [2x] 428 435 [2x] 438 442 446 472 479 481 [3x] 494 506 542 544 554 555 565 568 575 584 603)

This analysis gives striking confirmation to Pearsall's assertion of the pre-eminence of Hg. Fifty-one errors in El amount to one error every thirteen lines. By comparison with Hg, El at times appears careless: casually omitting words, or needlessly transposing words or substituting one word for another. A telling instance is NP 481: in the course of one line, El makes three copying errors, converting probable archetypal Saue ye I herde neuere man so

synge (Hg and 10 other witnesses) to Saue you herde I neuere man yet synge (El alone). There is little evidence, in this catalogue of misreadings, of the deliberate and careful editing some scholars see in El. These look like nothing so much as the typical short change of scribal inattention.

This marked difference between Hg and El creates its own issues. Most scholars accept that both manuscripts are written by the one scribe, identified by Linne Mooney as Adam Pynkhurst. How then is it, that this scribe is so careful when writing Hg, and so much less careful when writing El? In itself, this casts doubt on the assertion that the two might have been copied close in time to each other (even, in parallel).

Two final observations are in order. First, one must keep the errors in El in perspective. There is still reading after reading where El agrees with Hg, against almost all other witnesses, in retaining a difficult Chaucerian reading misunderstood by almost every other scribe: see the stemmatic commentary on NP 29, and the readings listed at NP 108 and NP 170. It is only in comparison with Hg that El appears inferior. El is at least as superior to other manuscripts as Hg is superior to El.

Secondly, although this analysis confirms the relative excellence of Hg, it is not perfect. The scribe was human when he wrote El, and made mistakes: he was human when he wrote Hg, and I list above five instances where it appears he miscopied the archetype. In these five, it seems El does preserve the archetypal reading, so making the editor's task easy. However, it is quite possible for both Hg and El to be in error at the same point. It appears this happened in NP 358, where Hg and El both read the line without hir eke: both a and Cx2 (and hence, the manuscript here used by Caxton) include these words, necessary for a metrical line. Barbara Bordalejo in her work on this Cx2 manuscript has identified at least two other points in the *Tales* where Hg and El are both in error, and one may identify the archetypal reading from other witnesses (particularly Cx2): RE 9 and CLT1065. Readers should always be aware of this possibility and keep in mind the motto ‘Trust nothing’.

5. Conclusions about the textual tradition

I here summarize the results of the analysis of tradition of the Nun's Priest and assocated links.

This analysis suggests that the majority of the witnesses represent three lines of descent:

> Through a 7 witnesses: a – Dd Cn Ma En1 Ds1 Ln Me;
> Through b 8 witnesses – Cx1 Tc2 Ii He Ne Cx2 Pn Wy
> Through γ 26 witnesses: c – Cp La Sl2; e – Bo1 Ph2; d1 – Ra2 Fi To1 ; d2 – En2 Ry2 Mm Ha2 Ph3 Pw Sl1; cd – Ry1 Dl Ht Lc Mg Mc Ra1 Tc1 Bw Ra3 Gl

In addition, analysis showed a group of eleven manuscripts (the o witnesses) which represent an uncertain number of individual lines of descent. These are:

> Hg Ch El Ad3 Ha3 Gg Py; the pairs Ha4/Ld1 Ad1/En3.

This latter group may represent as many as 9 separate lines of descent. In fact, it is likely that several of these beside Ad1/En3 and Ha4/Ld1 have shared exemplars but these copies may have introduced so few errors that it is not possible to distinguish the shared exemplars.

This leaves three witnesses as unclassifiable, Se Nl Ps. These show evidence of heavy editing and contamination, so can not be assigned to any group.

Finally, one must use this analysis with caution. The overall group divisions are clearly marked, as is the pre-eminence of a core set of witnesses (Hg El Ch, with Ad3 Dd Ha4 and, on occasions, readings from Cp and Cx2) as typically preserving the earliest forms of the text. However, many details remain (and may always remain) obscure. The exact relations of manuscripts closest to the presumed and lost archetype are uncertain, as are the groupings within the large and rather amorphous set of manuscripts descended from γ. However, the general consistency of the picture of relations emergent from the sections of text so far analyzed by The Canterbury Tales Project suggests that we are dealing with a single manuscript tradition, descended from a single copy (albeit of ambiguous status), not with many traditions descended from many separate exemplars arising

from sequential and distinct acts of publication.

6. Notes

1. For this ‘absolutist’ approach, see Kane’s preface to his edition of the A Version of *Piers Plowman*, 1960.
2. Other lines identified in our collation as additional to our collation base, and so likely to have been added by later scribes or revisers to the archetypal text, are: NP 61-1 61-2 62-1 66-1 101-1 106-1 111-1 114-1 127-1 218-1 471-1 474-1 474-2.
3. This theory of the nature of the originals – colloquially, the Ômessy deskÕ theory – appears to have been first proposed by J. S. P. Tatlock, most fully in his article ‘The Canterbury Tales in 1400’ 1935.
4. Compare too the retention of likely archetypal his in Cp Sl2, and its replacement by this in later γ manuscripts, in NP 157. See the stemmatic commentary ad loc.
5. In fact, Blake's position is actually considerably more subtle than this caricature suggests. See ****