The Witness Relations in Link 1 and The Miller's Tale

Peter Robinson

Introduction

Link 1 and The Miller's Tale are extant in 54 manuscripts and four print editions dating from before 1500. Of these 58 witnesses, four (Ad2 Hk Ox1 Ra2) lack more than 10% of the text, while the other 54 are relatively complete. This section presents our analysis of the textual relationships among these witnesses, based on an analysis of our collation, using the various tools available to us.

1. Phylogenetic analysis of the data

Following the completion of the transcription and collation processes described earlier in this CD-ROM, a NEXUS file containing a complete record of all agreements and disagreements among the 58 witnesses was generated from the parallel-segmentation collation apparatus. The NEXUS file format is widely used by evolutionary biologists to hold data concerning agreements and disagreements among populations of objects ('taxa' in evolutionary biology; 'witnesses' to us) at precise points ('characters' in their teams, 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

This line represents the variants at the word 'thus' of line 1 of Link 1: so, the label 'GL1L1_thus'. 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 of objects 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.

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.

The single NEXUS file, containing the full data for both Link 1 and for The Miller's Tale, was first processed six different times by PAUP, with each process producing a distinct set of trees, as follows:

1. Neighbour joining for each of three sets of data: for both Link 1 and The Miller's Tale separately, and for the whole of Link 1 and The Miller's Tale together
2. Maximum parsimony for each of the same three sets of data.

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

L1MI.nex	*file containing the full nexus format information for L1MI*
L1MI2nex.xml	*file containing all the XML apparatus from which these analyses derive*
Cx2-alpha	*folder containing parsimony analyses relating to Cx2 and the putative alpha group of witnesses*
Pars-L1MI	*folder containing parsimony analysis for The Miller's Tale and Link 1 together*
Pars-MI	*folder containing parsimony analysis for The Miller's Tale*

Pars-L1	*folder containing parsimony analysis for Link 1*
NJ-L1	*folder containing neighour joining analysis for Link 1*
NJ-L1M1	*folder containing neighbour joining analysis for The Miller's Tale and Link 1 together*
NJ-MI	*folder containing neighbour joining analysis for The Miller's Tale*
vargroupprofiles.html	*file containing the full variant group profiles for both Link 1 and The Miller's Tale. See below.*

Each folder contains the trees produced by the analysis. Each tree was saved in two forms:

1. 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
2. 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.

In this initial processing, the parsimony analysis was run through ten replications for each of the three sets of data. Briefly, 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 'shortest', or most parsimonious, tree or trees 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. This can be quite time-consuming, and so initially this was limited to ten replications. We found that for the Link 1 data, many trees of identical length were indeed found in the analysis, and so each replication took a considerable time. However, the analyses of the whole of The Miller's Tale and also for the tale and link together very rapidly found a single most parsimonious tree at each replication. This suggested that the results for Link 1 were due to there being insufficient data in the collation of these 76 lines; were it otherwise, we would have found the same pattern in analyses of the longer sections.

The speed of each replication for the longer sections suggested that we could increase the number of replications. Accordingly, we re-ran the parsimony analysis for both

The Miller's Tale and for Link 1 and The Miller's Tale together, for over a hundred replications each time. These yielded the trees labeled '-100' in the Pars-L1MI and Pars-MI folders. In fact, the 100-replications analysis was run several times more on each set of data: each analysis revealed exactly the same tree for each section. We can be reasonably confident that each tree gives the best representation of the data available by this method.

Nearly identical were trees found for each of the sections, that is for The Miller's Tale (MI) on its own as well as for Link 1 and The Miller's Tale together (L1MI). The trees can be found in the relevant folders. The only substantive differences between them are as follows:

1. In L1MI Py is rooted much closer to El Ch Hg than it it is in MI
2. In L1MI Hk branches closer to the a and b groups than it does in MI

Overall, the similarity of the two trees shows the consistency of the relations across the whole text. The memberships of the a b c d1 d2 e and g groups are identical across two trees, and the internal relationships within these groups are also nearly identical, where they are not indeed so.

The discussion hereafter is based on the trees created by the parsimony analysis for Link 1 and The Miller's Tale together: the L1MI trees. There are two trees: an unrooted cladogram and an unrooted phylogram.

2. The fundamental witness groups in Link 1 and The Miller'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.

2.1 Identifying the groups

Our initial hypotheses concerning the fundamental witness groups are based on the L1MI trees created by parsimony analysis.

It appears from the L1MI trees that the fundamental witness groups are as follows:

a: The pairs Cn Ma and En1 Ds1 with Dd
b: The groups Cx1 Tc2 Ne He Ox1 and Cx2 Wy Pn, and possibly Hk
c: Cp La Sl2 Ra3 and the pair Ln Tc1
d1: Dl Ha3 Ra2 Ht Nl, and possibly Se
d2: The pair Lc Mg with En2 Pw Ry1 Ld1 Ha2 Sl Ry2 Ld2 Gl Mm
e: The pair Bo1 Ph2 with Ra1 Bw Ad2
g: The pairs Ha4 Ii and Ad1 En3 (possibly with Gg Ps)

In addition to these the following manuscripts appear grouped in the L1MI trees:

o: the trio Hg El Ch with the pair Ha5 Ad3 and To1 and possibly Py Hk

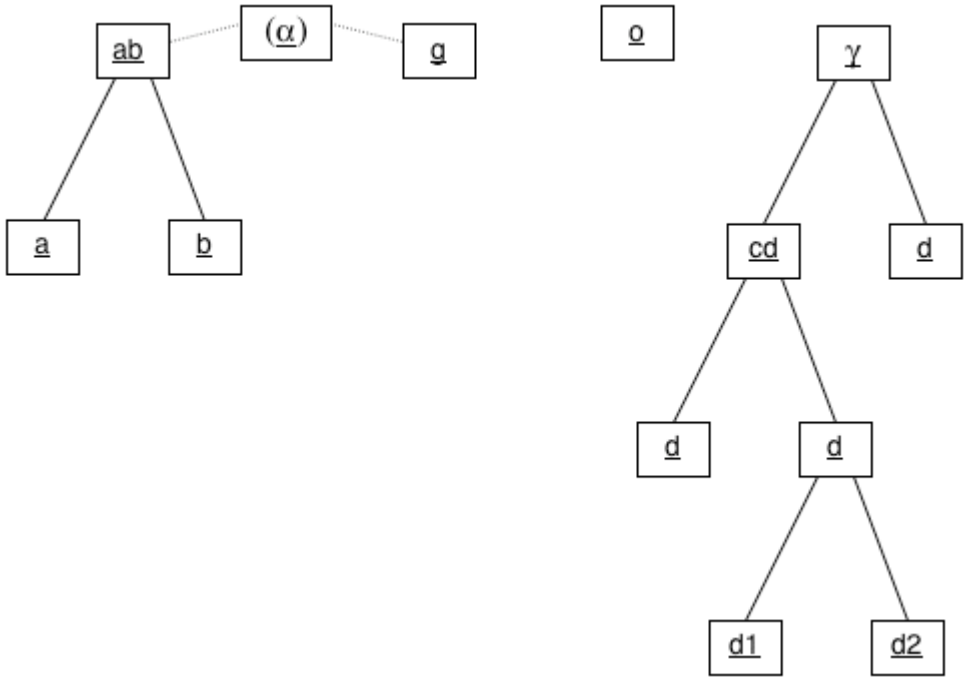
However, as the discussion below indicates, there is reason to think that while this last is shown as a group within the tree, the manuscripts within the group may not have a common ancestor below the archetype. To put it another way: their only common ancestor may be the archetype itself. Hence, their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

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 and Cx2. The length of this branch, including the section up to the rooting point of Hk, corresponds to around 200. As we will see later, variant database analysis suggests that there are some 221 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 branch to the common ancestor of the c witnesses, that is, to the node uniting Cp La Sl2 with Ra3 Ln Tc1, alongside the gamma symbol in the phylogram, suggests that there are rather few variants joining all the witnesses below this node. Indeed, variant database analysis suggests that there are around thirty readings introduced by the common ancestor of c. 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, as similarly En1 may be copied from Ds1.

Close scrutiny of the L1MI trees suggests the following relations among the fundamental groups:

1. It appears that a and b descend from a common ancestor, the ab ancestor
2. It may be (though this is doubtful; see the discussion below) that g may have a common ancestor with ab. The approximate position of this hypothetical ancestor is marked with α on the L1MI trees
3. It appears that d1 and d2 share an ancestor, the d ancestor
4. It appears that c and d share an ancestor, the cd ancestor
5. It appears that e and cd may share an ancestor. The approximate position of this ancestor is marked with γ on the L1MI trees

These relations may be summed up as follows. This divides the manuscripts into four, or possibly three, groups:



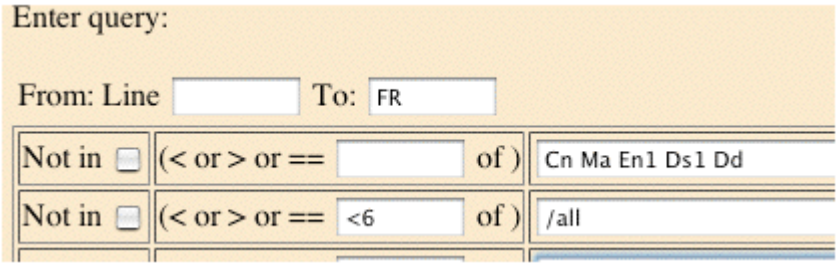
This arrangement suggests that we may be dealing with twelve different witness groups: e α ab a b γ cd e c d d1 d2. In addition, we are interested in the alignment of all manuscripts against the single most important division within the tradition, that between El and Hg. Accordingly we should like to 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 'hypearchetype'). We use the variant database facility, or 'VBase', built into the CD-ROM 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 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 five witnesses and nowhere else. Indeed, you can test this for yourself by running the following query in VBase:



This search produces just six hits: the replacement of ‘This’ by ‘The’ in the phrase ‘This sely Ialous housbonde’ in MI 218, of ‘herde’ by ‘seide’ in MI 253, the error ‘underpore’ in MI 279, and so on. But six 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 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:

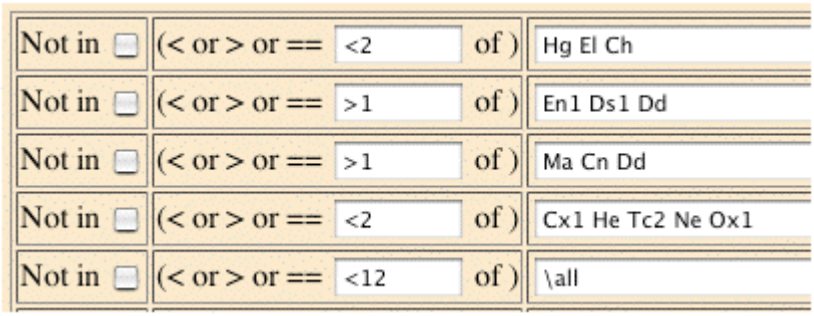
1. The variant was probably not present in the ancestor of the whole tradition.
2. The variant is present in a significant number of witnesses within the hypothetical grouping
3. The variant is not present in a significant number of witnesses outside the hypothetical grouping, and especially not in witnesses 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 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 "Fundamental 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 and two pairs, Ma Cn and 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 two pairs Cn Ma or En1 Ds1. It should be present either in both manuscripts of one pair, or in one manuscript of one pair and Dd, and also present either in both manuscripts of the other pair, or in one manuscript of the other pair 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 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 31 hits.

3. Assessing the variant groups

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

- The thirteen groups e α g ab a b γ cd e c d d1 d2 corresponding to groups apparently revealed by the MIL1 trees
- 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

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:

> L1 12 The Millere that for dronken was **al pale** (45 witnesses)
> a: The Millere that for dronken was **and** pale (10 mss including 4 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 4 of the a group suggests this is not an accidental agreement in a - as appear to be the other 6 instances of this reading, scattered throughout the tree (d2 b c and three d1 witnesses).

> MI 145 And pleyen **songes** on a small rubible (51 witnesses)
> a: And pleyen **song** on a small rubible (6 witnesses)

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

> MI 640 That **aswowne** lay bothe pale and wan (28 witnesses; **yet aswowne** 8

witnesses)

a: That **in swownyng** lay bothe pale and wan

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

Analysis of the variants in eleven of the groups suggest that the groups arose below the ancestor, through non-authorial copying. The eleven groups are ab g a b γ cd e c d d1 d2.

3.1 Lines present in El and not elsewhere

While many variants patently appear first within an exemplar below the ancestor **O** and are then mostly propagated by copies descended from **O** (thus, the variants referred to in the last section), there are two notable sets of variants in Link 1 and The Miller's Tale which do not conform to this pattern. They are the two pairs of lines, 46-1 and 46-2 from Link 1 in our numbering and 534-1 and 534-2 from the Tale in our numbering. The first pair is in the o witnesses El Ad3 Ha5 To1, in g, and in two d2 witnesses Ht Nl:

And euer a thousand ayeyns oon badde
That knowestou wel thyself but if thow be madde

The second pair is present in El alone of the o manuscripts, in b and a few manuscripts from d2 (Ha2 Ld1 Ry1)

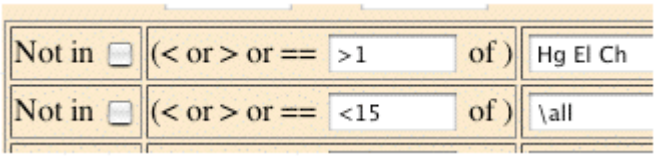
And vnto Nicholas she sayde stille
Now pees and thou shalt laughen al thy fille

One would be reluctant to assert that Chaucer could not have written these lines - the second pair has the unmistakable stamp of Chaucerian dialogue. The only witness which has all four lines is El. Elsewhere, I have argued that the first copyists of the text adopted different policies when faced with lines marked for addition and deletion, sometimes including them, sometimes excluding them. The El scribe, when copying El, seems to have followed a policy of including all the text he could, and so here; when copying Hg, the same scribe seems to have excluded such passages. Others seemed to follow no consistent policy, and this leads to the uneven pattern of distribution we see here.

Thus, special considerations are in play in the attestation of these four lines, and should not affect conclusions drawn from analysis of the whole body of variants in each group.

3.2 Groups which are not groups: o

Two groups of variants among those surveyed do not appear to pass the tests described above, and so may not descend from a single ancestor below the archetype. These are o and α . The test for o was rather simple:



This yielded 37 hits. This is comparable to the number of variants found for a and, on the face of it, are sufficient to constitute a group. However, analysis of the variants themselves tells a different story. Many of them appear of authorial character. The most striking of these in MI 605:

MI 605 I am thyn Absolon **my** dereling (o 6 witnesses; 2 'O my')
I am thyn Absolon **thyn** dereling (46 witnesses, with variants)

The shift within the direct speech, where o has Absolon first declare himself ('thyn Absolon') and then move to a parenthetic exclamation addressed to Alison ('my derelyng') is dramatic and sudden. We can imagine it working superbly in a live performance or reading. But it is exactly this shift which a scribe, working from a written exemplar, might fail to catch: and the evidence is that apart from witnesses close to the original (the trio El Hg Ch; but also To1 Gg Ps with the pair Ii Ha4 having the related 'O my'), every other copy failed to register this, and substituted 'thyn' for 'my' following the 'thyn' earlier in the line. Once this change was made, it was very unlikely to be reversed, and hence the complete absence of 'my' from elsewhere in the tradition.

Similarly, though less dramatically: o preserves a harder, more unusual reading, widely replaced elsewhere by a more commonplace reading in L1 4 (to->in), L1 31 (that I->I), L1 32 (preye->yow preye), L1 40 (fame->name), L1 56 (enquere->to enquere), MI 6 (leere->lerne), MI 11 (shal->sholde), MI 42 (man->men), MI 91 (ich->I), MI 132 (window->wyndowes, MI 265 and 271 (Astromye->Astronomye), MI 511 and 600 (cogheth->coghed), MI 641 (he brosten hadde->various replacements). Taken together, the concentration of these readings in relatively few witnesses close to each other at the centre of the L1MI 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.

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 Dd Cp La Bo2 Ad3 En1 Gg Ha4 Ha5 To1 Hk Py) argues in turn the closeness of these witnesses to the archetype of the whole tradition.

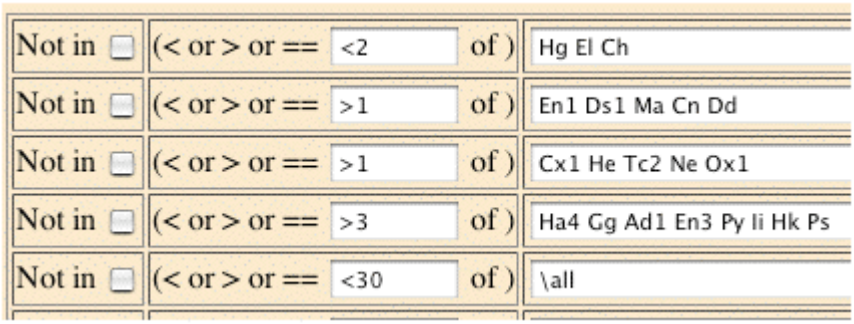
It follows that, as the o variants are ancestral to the whole tradition, their occurrence in these or any other witnesses does not suggest they have a common ancestor below the archetype of the tradition. These witnesses are of exceptional interest, because of the likelihood that at many points they 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 ancestral variants clustered in these witnesses. But one should not speak of an o group of witnesses; strictly, there is no such thing.

3.3 Groups which are not groups: α

The L1MI trees suggest that group g, and with it Ps Gg, shares an ancestor below the original with the ab group. Two further witnesses are rooted on the tree close to these, Py and Hk, and perhaps might belong with these. These witnesses include two which have long attracted (and puzzled) editors, Ha4 and Gg. That puzzle and the presence of this group close to the o witnesses mean that if this whole group does share an ancestor below the original, then it will be a group of the first importance. This hypothetical ancestor we may call α; it would be rooted next to the α shown on the L1MI trees.

The importance of this potential group is magnified by the placement of these witnesses on these trees relative to the other groups. Firstly, the apparent linkage of the key witnesses Ha4 and Gg with the ab group would be in accordance with these all having tale orders closely related to (if not identical with) the ‘Ellesmere’ or the type a order. In the stemmatic analysis on the General Prologue CD-ROM I suggested the existence of such an α ancestor, from which several of these witnesses and the ab a b witnesses all descended.

Early work identifying likely α variants and surveying their distribution across the witnesses suggested that the importance of α might be even greater than this. First, I used this VBase search to identify possible α variants:



To my surprise, it appeared that significant numbers of the variants identified by this search appeared in witnesses where I did not expect them: notably in the pair Bo1 Ph2, both of which seemed to contain around half the variants putatively characteristic of α . But it was clear from other evidence that Bo1 and Ph2 were descended from γ, the common ancestor of the large grouping of manuscripts containing the c cd d1 d2 e witnesses. It is an axiom that a witness cannot belong to mutually exclusive groups; so how could Bo1 Ph2 be members of both the α and γ groups? One explanation was that the whole γ group might descend from α. As this would account for all but a handful of all the witnesses of the tradition - all but the half-dozen or so o witnesses - this would indeed be remarkable.

However, the possibility that all the witnesses except the o witnesses might descend from a single ancestor below the archetype conflicts with other evidence. Firstly, it conflicts with the tale order evidence. As I observed above, the witnesses putatively descended from α all have the so-called type a order; none of the witnesses in the γ line of descent have this order. Secondly, it conflicted with the evidence provided by other searches. If there really was an α and it really was the ancestor for the whole large γ group (i.e. the c/cd/e groups), then one would expect to see these α variants in other witnesses within the γ group. But the distribution of these 19 variants within the group is very erratic. Among the key c witnesses, closest to γ we find only 1 in Cp, 3 in La, though 7 in Sl2: yet the more distant Ln has 10, Tc1 8. There are high numbers of these variants in the d1 grouping (10 in all of Fi Ha3 Nl Dl) but typically many fewer in d2, with the exception of Ld2 which has 8.

This odd distribution makes it difficult to maintain that the variants in this group have reached these witnesses by a process of introduction into a single ancestor followed by descent into these copies. This distribution is indeed rather like that seen in the o variants, where a variant is ancestral to the whole tradition (and so might appear anywhere) but is liable to be removed at any point (and so the different o variants have different patterns of distribution).

Is it that these variants are archetypal and were indeed present in the common ancestor, even though they are rarely present among the o witnesses (Hg 0, Ch 0, El 2, Ad4, Ha5 2, Bo2 5, To1 5)? But the character of these variants is not such as to

suggest they are authorial: they lack the distinctive character of those examined above.

In fact, these variants have quite the opposite character of the o variants. Those were typically the 'harder' reading, and highly likely to be removed, and most unlikely to arise by independent scribal action. These are typically the 'easier' reading; they are most likely to persist once introduced, and - the key point - they are highly likely to be introduced independently by different scribes at different points in the tradition. A few examples:

> MI 19 Whan **that** men sholde haue droghte or ellis shoures (o; 27 witnesses)
>
> Whan **[ ]** men sholde haue droghte or ellis shoures (ie, **that** not present: 22 witnesses: ab g c d1)

The pleonastic 'that' is characteristically used by Chaucer apparently for metrical reasons, and is characteristically apt for removal by scribes; so too MI 127, 517.

> MI 44 For youthe and Elde **is** often at debaat (o; 29 witnesses)
>
> For youthe and Elde **been** often at debaat (22 witnesses: ab d1 g; 3 others)

A classic substitution of the obvious 'been', giving agreement with the apparently plural subject, for the striking singular 'is'. Compare MI 142 (manere->maneres; MI 220 arm->armes)

> MI 61 She was **ful** moore blisful on to see (o; 25 witnesses)
>
> She was **[ ]** moore blisful on to see (ie, **ful** not present; 21 witnesses ab e d1 g; 1 other)

An easy omission of the additional intensive 'ful'.

From this, one has to conclude that there is little evidence that this is a distinct grouping, resulting from the introduction of these variants into a single ancestor. Rather, these variants are more likely the result of separate scribal actions, happening to cluster in the ab and g groups but actually introduced independently into each, and independently into other groups (especially the d1 e groups). Therefore they are not evidence either of a common ancestor α for the ab and g groups, or of a common ancestor for those groups and the γ group.

4. Difficulties with this analysis

In general, this analysis of The Miller's Tale gives strikingly consistent results across the whole tradition. However, one must understand the methods we use, and their limitations. Phylogenetic programs are, essentially, 'blind': they operate on data on agreements or disagreements, regardless of the source of that data or the reasons for

the agreements or disagreements. There are many safeguards, based on decades of experience, built into good phylogenetic software such as the PAUP program. These programs seek to discriminate agreements and disagreements resulting from genetic descent vertically from those which are caused by independent accident or are transmitted laterally. The programs are built on the assumption that the dominant model of variation and its transmission across the population surveyed is 'descent with modification', to use Darwin's phrase. In so far as this is true of textual traditions and of living organisms, these methods are sound.

However, there are notoriously within textual traditions (and indeed in living organisms) instances where this assumption is not valid. One can expect that in a tradition the size of that of The Miller's Tale, and with a text of its evident (and continuing) popularity, there will be cases of lateral transmission: scribes importing variants from other exemplars into their copies. The instance of the El and B witnesses, discussed below, seems such a case. One can expect too that such lateral transmission would have an effect on the trees produced by phylogenetic methods, as the tools attempt to accommodate the flow of variants laterally by distorting the trees of descent they hypothesize. The case of the variants introduced into Cx2 by Caxton and the effect this has on the relative rooting of the a b g groups and the o witnesses, discussed below, seems an instance of this.

4.1 Mixing of exemplars? El and the b witnesses

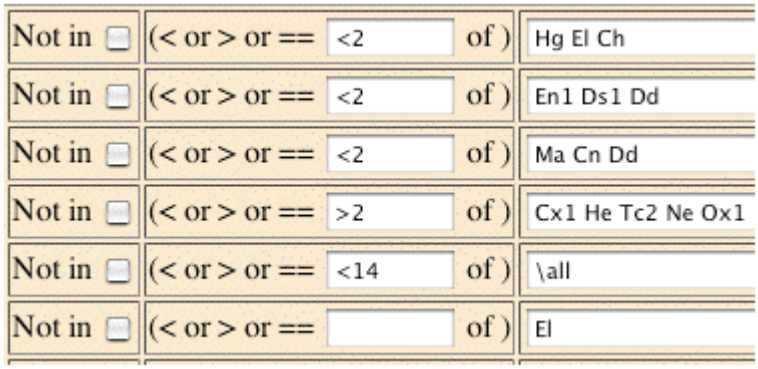
The L1MI trees show that El and the b group are distant from each other, in genetic terms. At least two exemplars - the ab and b exemplars - lie between El and any b witness. Accordingly, one would expect to find very few variants introduced by the b ancestor in El, or indeed in any o witness. The VBase search identified 222 likely b variants: you can see these by choosing 'Search VBase' from the pop-up menu in the centre of the toolbar, and then clicking on the 'B variants' link. The distribution of these variants in the o witnesses is as follows (you can gain these numbers by selecting the 'Count the hits in every manuscript' option in VBase after running the 'B variants' search, and them pressing the Submit button):

Ad3: 11
Bo2: 9
Ch: 3
El: 32
Ha5: 7
Hg: 0
To1: 27

Apart from El and To1, these numbers are consistent with random agreement: the

three in Ch, the nine in Bo2, for example, are there just by chance. But 32 (or 27, as in To1) seems too high a number to explain by chance. What are these 32?

We can run the VBase search as follows to see what they are:



We discover that fully fifteen of these 32 are in the two lines 534-1 and 534-2 (so too in To1). This pair of lines was discussed in 3.1 above, where I suggested that the transmission of these lines was subject to special factors and so not indicative of ordinary manuscript relations.

Removing these 15 reduces the number of b variants in El to 17: still high, but perhaps not too high to be explained by coincidence. Indeed, the character of some of the 17 is such that they could have arisen by coincidence: for example, the variants ‘on his gyterne’ (o ‘on a gyterne’) MI 147, ‘Til’ (o ‘And’) MI 257, ‘He seyde’ (o ‘And seyde’) MI 315, ‘sitten’ (o ‘seten’) MI 451, ‘a compaignye’ (o ‘compaignye’) MI 474. In a few cases, it appears that the reading is in fact ancestral to the whole tradition: that is, it appears by descent from **O** in El and the b witnesses but has been lost elsewhere. Thus the following:

> MI 49 A ceynt she werde **ybarred** al of sylk (13 witnesses: El Ad3 Ha5 To1 from o; b; e; Ht)
>
> 49 A ceynt she werde **barred** al of sylk (38 witnesses)

Here the dominance of the first reading in the o witnesses, the nature of the variant itself, and its occurrence in e all suggest that this was the reading of **O**. So too, though less certainly, for the very similar ‘ydight’ (13 witnesses including El and b)/’dight’ (42 witnesses) in MI 19.

But others in this group are so distinctive that it is difficult to argue that they could have appeared at several different points in the tradition. The most notable of these is this:

> MI 65 Tasseled with **silk** and perled with latoun (45 witnesses including all o except El)

> Tasseled with **grene** and perled with latoun (9 witnesses: El b [excluding Cx2 Wy Pn] To1 Se Dl Ht)

The variant is so striking that it is unlikely that it has arisen independently. On the other hand, precisely because it is so striking, it might have been remembered by a scribe who then copied it into a witness from a distinct line of descent. This would explain its appearance in To1 and Se (with Dl Ht likely having it by descent within the d1 group).

Another variant that might have travelled across witnesses by memorial contamination is:

> MI 511 And ofte he **cogheth** with a semy sown (10 witnesses, including o except El)
>
> > And ofte he **knocketh** with a semy sown (6 witnesses: El b [excluding Cx2 Wy Pn] Ha3)

The range of variants at this point ('coude' ' toughtet' 'singeth' 'sange' 'couched' 'spak') suggests that scribes were likely to be as puzzled as a modern reader by what action might be accompanied by 'a semy soun'. Again, one can suppose that 'knocketh' could have been remembered and then imported into a lateral copy. One could also account for the vivid 'amydde' in MI 622, in place of the more neutral 'in' for the line 'And Nicholas in the ers he smoot' in the same way.

One should also consider the possibility that these last three variants (MI 65 511 622), and indeed the lines 534-1 and 534-2 found in El and b, might have come from an oral performance of the Tale. All are dramatic in nature, occurring at high points of the narrative, and so would lend themselves to a recitation which might persist in a scribe's memory.

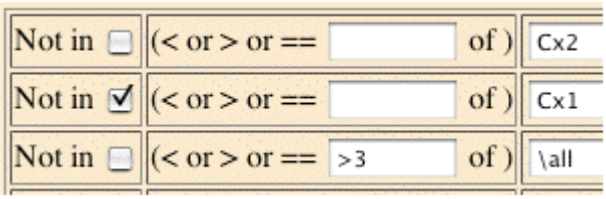
4.2 Caxton's corrections in Cx2, contamination and tree rooting

In the last section, I discussed two variants ('grene' MI 65 and 'knocketh' MI 511) present in all the b witnesses including Cx1, Caxton's first edition (here, with El) but not present in Cx2, Caxton's second edition, and its descendants Pn Wy. Analysis of the distribution of b variants shows too that Cx2 (again, with Pn Wy) has a rather lower number of b variants than other members of the b group. The quartet Cx1 He Ne Tc2 have between 163 and 218 of the 222 b variants (Cx1 Ne T all with more than 200). The trio Cx2 Pn Wy have respectively 132, 115 and 125 of the b variants.

The reason for this divergence between Cx2 and the other b witnesses is well-known. According to Caxton's own account in his 'Prohemye' to his second edition, after publishing his first edition some six years before, a 'gentylman' came to him, complaining of the quality of the text and declaring that his father had a better copy.

Caxton states that he ‘corrected my book’ by this copy. Analysis by several scholars, most recently by Barbara Bordalejo in her 2002 De Montfort University doctoral thesis, has shown that indeed Caxton did create his second edition not by resetting anew from this manuscript, but by writing corrections from the manuscript in a copy of his first edition, which the compositors then incorporated into the reset second edition. Though the extent of Caxton’s activity as corrector has been doubted by some scholars, Bordalejo showed that Caxton’s corrections were extensive and consistent, throughout the verse: according to her manual count, some 3000 substantive corrections in around 18000 lines or one every six lines.

This search locates all the differences between Cx2 and Cx1 and which are present in more than three witnesses (and so, not just in Wy Pn, which might inherit unique readings appearing in Cx2 by error, and not by correction from the better copy):



There are a total of 207 hits found by this search. While not all of these may have been imported from the better copy, study of their distribution in other witnesses suggests that the great majority were. This gives a correction rate of up to one correction every 3.5 lines, rather higher than usual on Bordalejo’s figures (though some of this difference may be the result of the use of computer methods here, as opposed to the manual methods used by Bordalejo).

A key issue in *Tales* textual scholarship is the nature of this better copy, used by Caxton as a source for many thousand readings in his second edition. It has long been recognized that this was a manuscript of high quality, a judgement confirmed by Bordalejo’s analysis. It seems that this manuscript had many readings known otherwise only from manuscripts long regarded as close to the archetype of the whole tradition; thus, Hg Ch El Ha4 Gg Ad3 and generally the manuscripts we have identified as o manuscripts. We can easily check this by studying the distribution of these 207 readings across the tradition. If the manuscript these (or most of these) came from was indeed close to the archetype of the whole tradition then we should see that the o witnesses will have a high proportion of these 207 readings. This is precisely what we find. The counts of these 207 for the o witnesses, and for other witnesses close to the archetype, are as follows:

Hg 188
Ch 188

El 183
Dd 182
Cp 174
Ha4 171
Gg 169

It must be expected that the presence of so large a body of variants - some 200 - appearing by lateral transfer in a single group of three witnesses must affect the workings of the phylogenetic programs. Indeed, there is some evidence that they did. In the folder 'Cx2-alpha' there are three phylograms made from the same L1MI nexus file, but for different combinations of witnesses, centring on the omission and inclusion of the Cx2 Pn Wy trio. The first of these, an unrooted phylogram, includes the trio, together with the central o witnesses Hg El Ch, the a witnesses Dd En1 Ds1 Cn Ma, the b witnesses He Ne Cx1 Tc2, and the g witnesses Ha4 Ii Ad1 En3 with Ad3 Ha5 Gg. This tree shows the following:

- the b group comes from a root quite close to the o witnesses (and indeed, closer to them than is any other group)
- the a and b groups are separated: they do not come directly from a common root but seem to be separated by the o witnesses
- b seems to share a common root with the g witnesses (with Ad3 Ha5 Gg)

A second tree takes exactly the same witnesses and the same data but excludes the trio Cx2 Pn Wy. The tree shows the same basic groups, but roots these very differently with respect to each other:

- the b group now shares a common root with the a group
- the a and b groups are not separated by the o witnesses
- b does not share a common root with the g group, and g is separated from both a and b by the o witnesses (with Ad3 Ha5 Gg)

This second tree agrees with the conclusion reached from analysis of the variant groups generated by VBase: that is, agreeing that there is a common ancestor for ab, and agreeing that g does not share an ancestor (putatively α) with ab below the archetype.

Thus, the first tree (including Cx2 Pn Wy) appears to give a distorted picture of the relations between a b g o; the second tree (excluding Cx2 Pn Wy) appears to give a picture more in accord with other analysis of the tradition. Thus, it appears that including C2 Pn Wy causes the phylogenetic analysis to distort the relative rooting points of a b g o. The reason is likely to be that the presence in Cx2 Pn Wy of a large number of variants otherwise typically found in o (and also in g) has led to this distortion. These variants actually moved to these witnesses by lateral transmission.

But the phylogenetic software seeks to account for variation by vertical transmission, not lateral transmission. This led to the software adjusting the relative rootings of a b g o, as it tried to account for the relatively high numbers of o variants in these three witnesses.

There are several lessons to be learnt from this exercise. First, it appears that while we can rely on the identification of clearly marked groups within the tree diagrams, we cannot rely on the rooting of these groups relative to each other. At the central points of the diagrams, where these rootings are to be found, the software is having to hypothesize relationships on increasingly uncertain data. One can see this in the relative shortness of the branches in the central parts of the phylograms. The second lesson is that we may expect contamination and other forms of methodical lateral transmission to influence the software to shift these higher-level rootings in ways which contradict other evidence. Because we suspected several other witnesses of contamination, I experimented with excluding these from the trees. This tree was made by excluding Se Py Hk, as well as Cx2 Pn Wy. It is interesting to compare this with the full trees including all witnesses.

These lessons mandate caution in the use of the tree data. One must use it as a starting point for analysis and use only conclusions which can be demonstrated from other analysis, especially from the VBase analysis of fundamental witness groups and their distribution.

5. Conclusions about the textual tradition

I here summarize the results of the analysis of the Miller and Link 1 tradition.

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

> Through ab 13 witnesses: a - Dd Cn Ma En1 Ds1; b - Cx1 Tc2 Ox1 He Ne Cx2 Pn Wy
> Through γ 30 witnesses: c - Cp La Sl2 Ra3 Ln Tc1; e -Bo1 Ph2 Ra1 Bw Ad2; d1 - Se Dl Fi Ha3 Ht Ra2 Nl; d2 - En2 Lc Mg Mm Gl Ld2 Sl1 Ry2 Ha2 Ld1 Pw Ry1

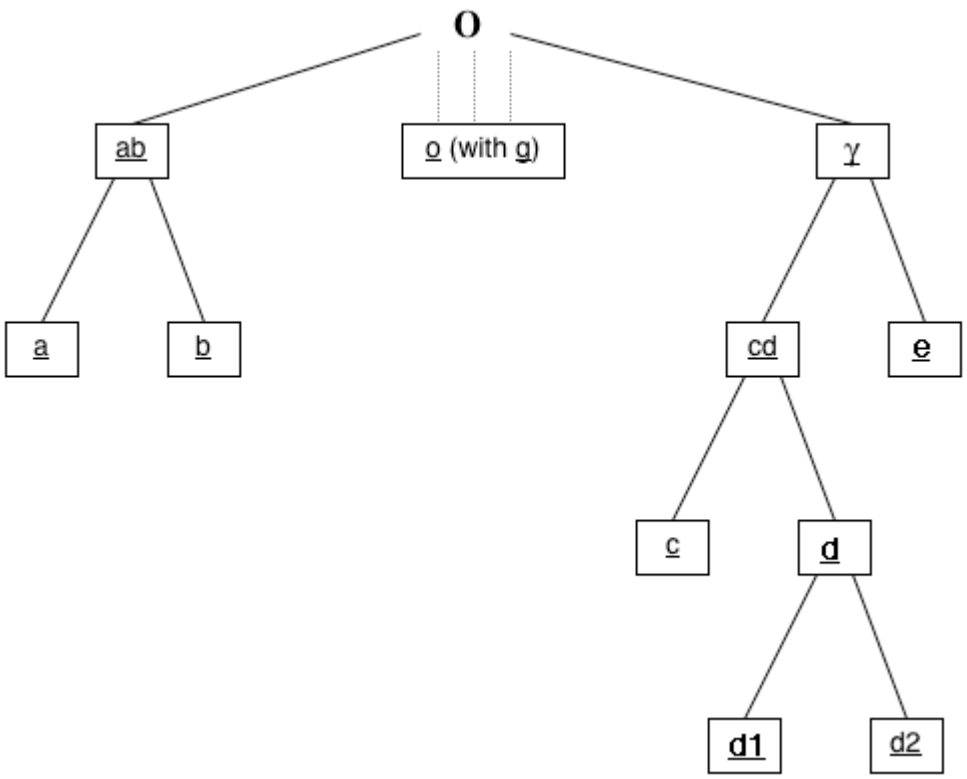
In addition, analysis showed a group (the o witnesses) which represents an uncertain number of individual lines of descent. These are:

> Hg Ch El To1 Bo2 Py Hk; the pair Ad3/Ha5; the quartet labelled g of Ha4 Ii Ad1 En3 possibly with Gg Ps

This latter group may represent as many as 11 separate lines of descent. In fact, it is likely that several of these beside Ad3/Ha5 and Ha4/Ii/Ad1/En3 have shared

exemplars but these copies may have introduced so few errors that it is not possible to distinguish the shared exemplars.

This may be represented schematically thus:



5.1 Comparison with the analyses for the Wife of Bath's Prologue and General Prologue

I here reproduce the summary discussion from the General Prologue CD-ROM;

> "The analysis of The Wife of Bath's Prologue suggested that the majority of the manuscripts represent just three lines of descent for the manuscripts, and so derive from three distinct copies of **O**:
>
> through α : El (second half); the exemplar of Cx2 ; hence, through ab all the manuscripts of the a and b groups
>
> through cd: Cp and all the manuscripts of the cd groups
>
> through ef: e (Bo1 Ph2 and others); f (Ld2 Ry2 Bw Ln)
>
> In addition, analysis showed a group (labelled o) of some seventeen manuscripts which could not be allocated to these groups, and so represented an uncertain number of additional lines of descent. These manuscripts were: Ad1 /En3 Ad3 /Ha5 Bo2 Ch El (also e) Gl (also c) Ha4 Hg Hk Ht Py Ps (also c) Ra2 Ra3 /Tc1
> The analysis in this workshop of The General Prologue tradition suggests, likewise, that the majority of manuscripts represent just three lines of descent, and so derive from three distinct copies of **O**:

through α : Ad1 Ad3 En3 Tc1; hence, through ab Ht Py Ra2 Ry1 ; and thence all the manuscripts of the a and b groups
through cd: Cp and all the manuscripts of the cd groups
through e: e (Bo1 Ph2)

In addition, analysis showed a group of some ten manuscripts (leaving aside the fragmentary Do) which could not be allocated to these groups, and so represented an uncertain number of additional lines of descent. These manuscripts were: Bo2 Ch El Gg Ha4 Hg Ln Ps Ra3 To1. Seven of these ten (all but Gg Ln To1) are among the seventeen classified as o group in The Wife of Bath's Prologue. Within this ten, it appears that El and the trio Ch/Ha4 /Hg may represent two independent lines of descent."

The analysis of The Miller's Tale differs from these in the following details:

1. Evidence for α is lacking. Thus, witnesses previously seen as α are here assigned to the o witnesses: thus Ad1 Ad3 (El and Cx2 in part of WBP)
2. The e group appears here to descend from γ, the common ancestor of cd. The group identified as f in WBP is here merged with e
3. Several witnesses move between groups in different tales. Thus:

    1. Gg is e in WBP, o in GP MI;
    2. Gl is d and o in WBP (with shift of exemplar); d in GP MI;
    3. To1 is d in WBP, o in GP MI;
    4. Ln is e (f) in WBP, o in GP, c in MI;
    5. Ld2 Ry2 are e (f) in WBP, d in GP MI;
    6. Ry1 is apparently ab in GP, D in WBP MI; Nl is b in GP, d in WBP MI;
    7. Ra3 Tc1 are o in WBP, c in MI, while Ra3 is c in GP, Tc1 is α in GP.

In summary, we find in all three traditions the following:

- Most witnesses descend through two lines of descent, essentially the ab and cd groups identified by Manly and Rickert. In GP WBP there is a third line, e (ef in WBP); this merges with cd in MI
- Some fifteen witnesses in all three cannot be assigned to these two lines of descent and form an uncertain number of lines of descent. In all three, these include Hg El Ch Bo2. These are the witnesses we call o.

I argued in the "Stemmatic Analysis" section of the General Prologue CD-ROM that the similarities between the three traditions are most easily explained by their arising from a single set of exemplars through uniform acts of copying, and not by there being separate texts of the separate parts of the Tales, each with its own history. The differences between the traditions are, I argue, such as might arise in the course of the copying of a long text divided into many parts, where a scribe may move from one

exemplar to another between sections (and sometimes within sections). Some other differences may be apparent rather than real. It is possible that e is always a relative of the cd group, while the evidence for α is tentative at any point.

5.2 Comparison with the conclusions of Manly and Rickert

Manly and Rickert give their account of Link 1 and the Miller's Tale on pages 136-154 of their volume II. A major difference between this account and theirs is that they see a shift of relationships around A 3480: MI 294 in our numbering. They argue (II 139) that Bo2 and Ad3 join a at that point, while Ad3 is with Gg up to that point, and Ra2 and Nl join the group we call e at that point. There are just four variants which support a relationship of a Bo2 Ad3 after line 294: in 296 392 467 604. This seems rather few. There is only one variant supporting a link between Nl and e after 294 (in 316) and none supporting a link between Ra2 and e.

A second major difference is that Manly and Rickert argue that the groups we label g b and Ad3/Ha5 are linked throughout. They cite nine variants in support of this contention. Three of these, in MI 61 (=A 3247) MI 300 (=3486) and MI 457 (=3643), are among the group of variants we examined for evidence of an α group. We concluded that these variants could not be used to support the existence of any group below the ancestor, and so could not be used to support a common ancestor here for b g. The other six are:

> MI 99 (=3285 'quod she': a manifestly easy substitution for o 'quod ich' ; cf. 607 for the confusion of pronouns in direct speech)
> MI 141 (=3327 'and' for o 'or'; again an easy substitution)
> MI 210 (=3396; the variant pattern here is so complex that little can be drawn from it);
> MI 373 (=3559 'wit' for o 'wittes')
> MI 386 (=3372 'Into' for o 'Vnto');
> MI 585 (=3773 ''this' for o 'his').

None of these carry conviction. In several of these, the variants are in 25 or more witnesses, usually a sign that the variant is likely to arise independently at various points. This is likely to be the explanation here.

As elsewhere, Manly and Rickert's analysis is sound in identification of the basic groups. However, their method has difficulties when trying to discover links between groups closer to the archetype. They have particular difficulties with the witnesses we identify as o: they seek to fit these into groups on necessarily scanty grounds. An instance is their handling of Bo2: as explained above, their attempt to align this with a after MI 294 is based precariously on just four variants.

6. Provisional classification of the manuscripts by their Variant Group profiles

Following the identification of the fundamental groups discussed above, I created a 'Variant Group profile' for each witness. This gives the number of variants from each of the fifteen groupings in each witness. The number is further broken down by section: Link 1. MI 1-150, 150-300, 300-450, 450-end. Analysis of the numbers of variants in each section might reveal cases of shift of exemplar.

For incomplete manuscripts, I also give in square brackets a scaled figure for the number of readings in the group. The total number of readings in the manuscript gives some idea of how fragmentary the manuscript is. A full manuscript has around 6000 readings. Thus, Ad2 with 3368 readings has about 55% of the text. Therefore, the 17 e readings found in Ad2 scales to 33, adjusting the number as if the manuscript had a full text, and I express this as follows: 17 [33] of 37.

Ad1	g (115 of 117); the cd (6 of 15) and γ (9 of 26) variants appear probably by coincidental agreement; cf En3. For the α variants here see above.
Ad2	e (17 [33] of 37) and hence γ (12 [23] of 26). No other affiliations.
Ad3	o (11 of 37). 15 of the 26 g readings are for the two lines L1 46-1 and 46-2 (see 3.1); the other 11 then will probably be by coincidental agreement
Bo1	e (34 of 37); and hence γ (17 of 26)
Bo2	o (19 of 37); no other affiliations. There is no support for Manly and Rickert's hypothesis that Bo2 joins b after MI 294: it has only 1 (of 103) b variants for MI 300-end. Bo2 tends to agree with Hg against El (75 of 120 agreements with Hg as against 29 of 149 with El)
Bw	e (17 of 37); and hence γ (24 of 26). Some contamination with cd is possible, as Bw has 13 of 15 cd variants; but the distribution of variants between cd and γ is uncertain
Ch	o (31 of 37); no other affiliations. Ch tends to agree with Hg against El (83 of 120 agreements with Hg as against 34 of 149 with El). I suggested in the General Prologue analysis that Ch might share an exemplar with Hg
Cn	a (31 of 31) and hence ab (15 of 17)
Cp	c (12 of 14) and hence cd (13 of 15) and γ (22 of 26)
Cx1	b (218 of 222) and hence ab (13 of 17)
Cx2	b (132 of 222) and hence ab (7 of 17). See 4.2 above
Dd	a (31 of 31) and hence ab (14 of 17). Dd tends to agree with Hg against El (75 of 120 agreements with Hg as against 37 of 149 with El)
Dl	cd (11 of 15) and γ (15 of 26). Compare En2
Ds1	a (30 of 31) and hence ab (15 of 17)
El	o (29 of 37). See 4.1 for discussion of the 32 b variants in El
En1	a (31 of 31) and hence ab (16 of 17)
En2	d (20 of 33); cd (13 of 15) and γ (23 of 26). Compare Dl.

En3	g (115 of 117). The cd (5 of 15) and γ (8 of 26) variants appear probably by coincidental agreement; cf. Ad1. For the α variants here see 3.3
Fi	d1 (24 of 42), hence d (28 of 33); cd (14 of 15) and γ (19 of 26)
Gg	Possibly g (32 of 117); no other affiliations, and 10 (of 37) o variants suggest an o grouping
Gl	d2 (23 of 50), hence d (31 of 33); cd (13 of 15) and γ (18 of 26)
Ha2	d2 (24 of 50), hence d (32 of 33); cd (13 of 15) and γ (20 of 26)
Ha3	d1 (21 of 42), hence d (17 of 33); cd (12 of 15) and γ (10 of 26). This manuscript has the highest number of unique readings of any witness to L1MI: 384, or one every two lines
Ha4	g (96 of 117); no other affiliations, and 11 (of 37) o variants suggest an o grouping
Ha5	o (11 of 37). 15 of the 24 g readings are for the two lines L1 46-1 and 46-2 (see 3.1); the other 9 then will probably be by coincidental agreement
He	b (163 of 222) and hence ab (11 of 17)
Hg	o (32 of 37)
Hk	The 18 (of 222) b variants may arise from coincidental agreement, and there are no other affiliations. By default, it is then o, but its low count of o variants suggests there may be several intervening copies between it and **O**
Ht	d1 (30 of 42), hence d (30 of 33); cd (13 of 15) and γ (17 of 26)
Ii	g (104 of 117)
La	c (11 of 14), hence cd (12 of 15) and γ (23 of 26)
Lc	d (26 of 33), hence cd (12 of 15) and γ (16 of 26). Cf Mg
Ld1	d2 (31 of 50), hence d (30 of 33); cd (10 of 15) and γ (18 of 26)
Ld2	d2 (27 of 50), hence d (29 of 33); cd (13 of 15) and γ (22 of 26)
Ln	cd (10 of 15), hence γ (19 of 26); cf. Ra3 Tc1
Ma	a (27 of 31) and hence ab (17 of 17)
Mg	d (26 of 33), hence cd (12 of 15) and γ (17 of 26). Cf. Lc
Mm	d2 (34 of 50), hence d (32 of 33); cd (13 of 15) and γ (25 of 26)
Ne	b (217 of 222) and hence ab (11 of 17)
Nl	d1 (37 of 42), hence d (27 of 33); cd (11 of 15) and γ (10 of 26). Nl has the second highest number of unique readings of an witness: 288, or approaching one every two lines
Ox1	b (30 [110] of 222), hence ab (2 [8] of 17). Incomplete and so difficult to assess
Ph2	e (32 of 37); and hence γ (15 of 26)
Pn	b (115 of 222) and hence ab (6 of 17). See 4.2 above
Ps	g (36 of 117). This manuscript is much corrected, and the differing proportions of g variants within it (20 of 30 in L1; 16 of 87 in MI) may support this
Pw	d2 (28 of 50), hence d (30 of 33); cd (14 of 15) and γ (23 of 26)
Py	Most of the 23 g variants are contained in the two lines L1 46-1 and 46-2. Like Hk, it is o by default, but its low count of o variants suggests there may be several intervening copies between it and **O**
Ra1	e (25 of 37), hence γ (23 of 26)

Ra2	d1 (13 [37] of 42), hence d (27 of 33). The very few cd and c variants may result from its fragmentary state
Ra3	cd (10 of 15), hence γ (18 of 26); cf. Ln Tc1
Ry1	d2 (26 of 50), hence d (29 of 33); cd (15 of 15) and γ (23 of 26)
Ry2	d2 (46 of 50), hence d (33 of 33); cd (13 of 15) and γ (25 of 26)
Se	d (20 of 33), hence cd (11 of 15) and γ (13 of 26). This may be d2 within d (12 of 50). This manuscript elsewhere mixes exemplars and this may also be occurring here
Sl1	d2 (37 of 50), hence d (28 of 33); cd (14 of 15) and γ (23 of 26)
Sl2	c (10 of 14), hence cd (14 of 15) and γ (23 of 26)
Tc1	cd (10 of 15) and γ (18 of 26). Cf. Ln Ra3
Tc2	b (203 of 222) and hence ab (16 of 17). Possibly a copy of Cx1
To1	o (13 of 37). To1 tends to agree with El against Hg (50 of 120 agreements with Hg as against 82 of 149 with El). For the 27 (of 222) b readings in To1 see 4.1
Wy	b (125 of 221) and hence ab (7 of 17). See 4.2 above

7. Notes

1. For this ‘absolutist’ approach, see Kane’s preface to his edition of the A Version of *Piers Plowman*, 1960.