nature portfolio

Peer Review File

Genetic population structure across Brittany and the downstream Loire basin provides new insights on the demographic history of Western Europe

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Summary:
In this paper the authors investigate the fine-scale population structure of northwestern France using a combination of haplotype sharing methods (e.g. chromopainter and finestructure, and IBD analysis) and rare variant analysis incorporating both modern and ancient individuals. Their analysis reveals extensive subtle population structure in northern France, comprising 154 genetic clusters, which segregate geographically and at coarser levels (e.g. k=3 and k=18 for TVD tree) show relationships with the distribution of surnames and linguistic groupings across France. The authors posit that their observations support the hypothesis that language and rivers have played some role in shaping the genetic structure seen in northwestern France. In addition the authors explore changes in demography across time and ancestral contributions from Europe. Notably they discover putative evidence of a minor population contraction in Brittany aligning to the period of the Black death, and a strong ancestral link between Brittany and Ireland and West-Britain. Finally the study explores the relationship between modern and ancient individuals. They find striking evidence of genetic continuity in France from the Mediaeval period to present day and observe substantial steppe-related ancestry north of the Loire, with particularly high allele sharing between people from Western Brittany and individuals associated with the Bell Beaker culture.

I believe the reported analysis will be of interest to readers of Nature Communications as it demonstrates several important findings about the population history of northwestern France. The manuscript is well written and the majority of analyses appear robust, however I have a number of questions and suggestions that I would like to be addressed prior to publication.

Major comments:
- Correlation between clusters and dialectal areas (Lines 241-254 and S figures S1.11-S.1.12: I am unclear how you are relating your clusters to changes in dialectal areas in this section. As far as I can see there is nothing in the methods section describing what metric is used to correlate genetic clusters with dialectal areas, and the supplementary figures are left unexplained (no figure legends, unclear what exactly these are showing). To itemise my questions:
1.) Are you quantifying overlap between clusters and dialectal regions, or have you made a qualitative observation of overlap? (if the latter I think this should be stated in the text as it is currently presented as a result rather than an observation).
2.) What are the maps in S Figures 1.11 and 1.12 actually showing? Are these just maps of the dialectal regions or are they showing some measure of correlation with your clusters (The phrasing “in correlation with cluster [x]” in the figure titles suggests to me you have somehow measured correlation between dialectal area and genetic cluster, but the figures appear to be just regional maps outlining the dialectal areas).

If there is no formal quantification of the correlation between the dialectal regions and your clusters I

would suggest doing the following to allow the reader to assess the degree of correlation for themselves:
i.) Plot all the samples from just the relevant cluster (e.g. Bretagne-Centre for S1.11 and Cornouaille for S1.12) on the supplementary dialectal maps to show how well they overlap with the dialectal areas.
ii.) Report the percentage of samples in the relevant clusters falling within the main boundaries of the dialectal areas. (You could define a polygon covering the majority of the dialectal area and perform a point in polygon analysis to count the number of samples with geocodes falling in that region)

Another suggestion would be to formally test the relationship between the geographic region corresponding to the dialects and your genetic clusters, following the framework laid out in Supplementary note 9 in Bycroft et al. 2019 (https://www.nature.com/articles/s41467-018-08272-w#Sec6)

- Quantifying the effects of rivers on structure: The correlation between cluster borders at finer scales and rivers other than the Loire is interesting, but I wonder if the notion that rivers influence structure can be quantified a little more solidly than visual inspection of these boundaries:
i.) Could you test if neighbouring clusters on either side of a river typically show stronger Fst than neighbouring clusters not separated by a river (at a similar geographic distance)? (as you observed for the Loire at k=3 when considering Fst for EBP vs SLO compared to EBP vs WBR).
ii.) Alternatively could you fit a geo-genetic model such as EEMS (Petkova et al. 2016) or MAPS (Al-Asadi et al. 2019) and examine if there are lower estimates for migration across these rivers? The idea here is to determine if the differentiation between samples across a river is greater than expected by a pure isolation-by-distance model.
Either of these analyses would lend more weight to the statement that rivers other than the Loire influence structure.

- Further analysis of IBDNe bottleneck in EBP and SLO: Based on your simulations of different demographic scenarios you suggest that population structure could generate a similar IBDNe curve dip as you observe in EBP and SLO without a real bottleneck. Seeing as you have identified structure in these regions (clusters at finer finestructure and TVD-tree splits), it would be worthwhile to rerun IBDNe on the largest subcluster from these regions, which should be less structured (i.e. take the biggest subcluster within EBP at k=39 or k=154; assuming this is large enough to be run in IBDNe) and see if the trend remains? If it does this would give more confidence that the bottleneck is real (Although I note that your gene-flow simulation also produced a similar curve).

- Globetrotter analysis results incompletely reported: You state in the text that “all seven French populations showed evidence of admixture ($P<0.0001$)” in your GLOBETROTTER analysis, and describe in the methods section the process by which you infer confidence intervals for admixture dates etc, however these dates are not reported anywhere (unless I have missed something). Additionally I believe your p-value is actually $P<0.01$ (see ii. below):
i.) Could you provide some sort of a table (supplementary would be fine) summarising the key output of GLOBETROTTER for each population tested (e.g. “best-guess” conclusion for admixture, Major and minor admixture sources, confidence interval for event timing (generations or years)). The variation in dates and number of admixing sources may be of interest to readers.

ii.) I am unclear how you get P<0.0001 with only 100 bootstraps (100 bootstraps reported in your methods section). To quote the GLOBETROTTER manual “the proportion of inferred date(s) that are <= 1 or >=400 give you the p-value for any evidence of detectable admixture.” This means that with 100 bootstraps the lowest p value estimate you can get is P<0.01 (i.e. less than 1 in 100). To report a p<0.0001 you would need 10000 bootstraps…

Minor comments:
-Figure 1 and lines 172-180: You define the WBR, EBP and SLO clusters clearly in the text, but it may help readers to interpret the figure and results more readily if these clusters are also labelled in figure 1a (Simplest option is including a key in panel a mapping WBR to yellow, EBP to blue and SLO to red).

- What is meant by “levels of clustering” Line 165-167: I am unclear what is meant by the statement “At this finest scale, levels of clustering were similar to those previously found for Spain (3), and larger than those previously reported for France (31) and Great Britain (5).”
Do you mean the number of clusters was larger or the size of the clusters? If you mean the number of clusters could you comment on what your interpretation of this is and why it is important to report?

I would expect that the raw number of clusters found by finestructure is somewhat a function of the sample size, and not necessarily a good measure of the degree of structure in the population being analysed. For example, the previous study of France cited in this comparison uses samples from all regions of France, but has a much smaller sample size which likely explains the lower number of clusters (If sample sizes were equal we would probably expect to see more structure in the whole country than just the northwestern region).

-IBDNe simulations: In your simulation study you state “We ran the IBDNe software on our data only using IBD segments that were 4cM in size.” I think you mean “greater than 4cM” here

-Line 286 in main text referring to fig S1.15: I think you mean “minimum centimorgan length” here not “chromosome length”.

-Supplementary figure 1.14: This figure could be made clearer by increasing the width of the plot lines in the map for the borders and rivers. At present the thin lines make it hard to distinguish borders from rivers. In addition the colour scheme for clusters in panel three is quite hard to distinguish due to the use of similar shades of green being used. You might consider changing the colour scheme or alternating plot characters (i.e. circles, triangles etc) to make this figure clearer.

- Plots of finescale structure based on rare variants: In figure 2 have plotted your rare variant sharing clusters using picharts per département rather than plotting each individual on the map coloured by their cluster (as in finestructure map plots like figure 1 and S.1.1 etc). Is there a reason for this visualisation choice? I’d be curious to see the raw geographic distributions of the clusters but understand that this might be a neater visualisation given the likely overlapping cluster boundaries.

Signed:
Ross Patrick Byrne

Reviewer #2 (Remarks to the Author):

The manuscript titled "Genetic population structure across Brittany and the downstream Loire basin provides new insights on the demographic history of Western Europe" presents a comprehensive investigation into the genetic history of France. To do so, Alves, Giemza et al. generated novel genomic data for approximately 4,000 individuals from the northern half of France and analyzed it, together with publicly available data for present-day and ancient European individuals, with state-of-the-art methods in population genetics. The authors skillfully employ a detailed population genetics analysis and integrate it with linguistic data, geographical data, family names, and historical context to unravel the intricate population dynamics that have shaped the present-day genetic landscape of Northwestern France and describe the population structure of the region in detail.

The manuscript's strengths are notable right from its outset. The authors identify the critical role of Northwestern France as a geographical meeting point of three main ancestral European populations, namely Western Hunter-Gatherers, Early European Farmers, and the Yamnaya Eurasian Herders from the Steppe. This central positioning lends significance to the study of the genetic history of the region. The comprehensive dataset of approximately 4,000 present-day individuals from Northern France, along with the inclusion of ancient DNA data from six medieval individuals from Western France, highlights the comprehensiveness of the study's sampling strategy – another notable strength of this work. This approach not only allows for the exploration of fine-scale population structure (as the authors thoroughly examine in this manuscript) but also presents an opportunity to address the impact of historical migration events (including events during the Pre-History, Antiquity, and the Migration Period). I find it particularly interesting, for instance, the authors' insights about the relationship between Brittany and Ireland based on their genetic data.

In my point of view, all methods were used to the highest standards, and all conclusions are fully supported by the data. I have not identified any major weaknesses with this manuscript; however, I have a few minor comments below and would like some clarifications about the ancient DNA (aDNA) data and analyses. Knowing the previous works of some of the authors, I believe the aDNA analyses were correctly performed. I do, however, think that more details are needed. I apologize in advance if the issues I bring up are already explained in the manuscript and I have missed them. If so, I would appreciate it if the authors could point me to the relevant sections in the text.

In sum, I believe the take-home message from this paper is exciting and of broad interest. This work presents a technical tour de force and reports a broadly useful data resource. This paper, when published, will be of great interest to those in the fields of population genetics, anthropological genetics, human genetics, and genomics. The dataset generated by this work will also be an extremely valuable resource for these fields.

Thank you for the opportunity to review this excellent manuscript.

Best,

C. Eduardo Amorim

~~~~~~~~~~~~~~

COMMENTS ON THE aDNA ANALYSES

I have a few concerns related to the authentication of the aDNA data and whether the necessary filtering of the data was performed.

1. Were the aDNA data authenticated somehow? I see a table describing the depth of coverage and contamination rates; however, I did not find any mention as to whether post-mortem damage (PMD) patterns in the DNA sequencing reads were assessed. In my point of view, PMD patterns must be assessed in a study such as this. If the PMD patterns were assessed in the aDNA samples, are they typical of aDNA? I kindly request the authors to provide some information about this, perhaps a figure or additional columns to Table S2.1 describing the percentage of terminal bases with C-to-T and G-to-A transitions.

2. On a related note, how PMD was taken into account when calling variants? I'm assuming the samples were not UDG-treated. If so, were transitions included in the analyses? If not, why? Please consider mentioning whether UDG treatment was performed and how.

3. There is no mention of the high (>10%) contamination estimates for fra008 and fra017 (X-contamination column Table S2.1). Reading the manuscript, nothing strikes me as a bias introduced by contamination, but I wonder if the authors could perhaps reassess their conclusions in light of the potential biases introduced by contamination. Perhaps just a brief mention of it in the manuscript should suffice.

4. More details are needed in the aDNA-related methods description.

a. Was DNA extraction and library prep done in a clean aDNA-dedicated lab?

b. Please add a supplemental figure with the radiocarbon dating results.

c. Were the libraries single or double-stranded?

d. I believe paired-end reads were used, but I couldn't find any clear mention of it. (Forgive me if I missed it)

e. Were duplicates removed? The last paragraph of page 27 describes a methodology that I am unfortunately not familiar with. Consider rephrasing it to add more details.

f. Table S2.1 is a bit confusing. Are these dates based on radiocarbon dating or archeological context? The manuscript mentions the former but, to the best of my knowledge, the format is unusual for radiocarbon dates. Also, I think the column "HOA sites" is distracting because, I believe, it has the same
~~~~~~~~~~~~~~

value for every row – consider excluding this column and adding the of number sites to the table description. Last, is the “mean cov.” column describing the mean coverage in the HOA SNPs or else? – unclear to me.

5. One small editing suggestion I would like to make is that you move the aDNA methods part to somewhere closer to the beginning of the Methods section. Although I understand this is a matter of writing style and the authors may disagree with this, I believe this would make the text more logical to the reader. For instance, you say somewhere that you use “inbred:YES” for the analysis and that implies that the data is in the pseudo-haploid format. However, as far as I can recall, this is not mentioned anywhere before that point. It is mentioned later, on pages 27 and 28.

6. Last, but not least, please add the source of the ancient and present-day data you used. In my opinion, it is unfair to the authors of the papers from which you used data to not have their papers cited. Citing the AADR preprint or website is not enough (by the way I believe you should cite the preprint, not the website). I quote the preprint by Mallick et al. on bioRxiv here:

“Researchers who use the curated dataset from the AADR as the basis for analyses should cite this paper and the version of the AADR they downloaded, including a reference to that version’s doi. Citing the AADR paper is not a substitute for citing the original publications that produced data, which should be specifically referenced in each publication.”

I understand that Nature Communications might have a limitation on the number of citations that can be included in the main text. If that’s the case, to avoid going over this number, I suggest authors cite these papers in the supplemental material, therefore acknowledging the hard work of several people who contributed the excellent and very useful resource that AADR is.

MINOR COMMENTS:

[Table 1] What does “published” in the name of the samples on this table and elsewhere mean?

[Fig 4a] What are the numbers in the parenthesis after the name of the groups? Sample sizes, citations, or else?

[Page 20] The description of the SNP array genotype data generation is very detailed and mentions the exclusion of related individuals and the filtering of SNPs. The next section describes the generation of WGS data. Was relatedness also assessed within the WGS data? Were SNPs with minor allele frequency (MAF) <5% also removed from the WGS dataset? Have authors tested if there were batch effects introduced by different sequencing strategies (SNP capture vs. WGS)? Are these 856 WGS distributed evenly across the studied region? I apologize in advance if I missed this information.

[LD pruning] Was the dataset LD-pruned for fineSTRUCTURE (as it was done for the PCA)? If not, do authors believe it could introduce bias in their results? In my experience, fineSTRUCTURE is sensitive to how the SNP set is filtered. I would like to know whether the authors, who are more experienced than I

in this type of analysis, think this might have introduced any bias that may have affected their observations. Thank you!

[Page 21] It is unclear why authors highlight the Viking dataset from Margaryan et al. 2020 within the data included in the AADR. Is this because this data was not included in version 42.4 of AADR? The dataset is included in AADR currently – therefore my confusion.

[Page 22] Please provide more information about the analyses performed in R. Was a publicly available R library used or an in-house script? Please cite the former. Consider making the latter available for reproducibility and transparency, if applicable.

[Page 24 – RoH] This section might need to be more detailed and include the detailed parameters used for estimating RoH and IBD unless RefinedIBD has a standard setting of parameters. If so, please mention the default parameters that were used. I am not familiar with RefinedIBD but I have used other methods to run these analyses and observed that results may be sensitive to the choice of parameters such as missing data allowed, window size, number of heterozygote sites allowed within an RoH fragment, and minimum length of the fragment, among others. Therefore I belive you need to be more specific about what parameters you used here. Sorry if I missed this information from the text.

[Lines 953 – 955] This sentence seems to be incomplete.

[Data availability] Are authors planning to make their data available? If so, how? Please clarify.

[PCA Fig S4.1] Why did authors use the HOA instead of the 1240K for this and other analyses?

[Typos] Here are two potential typos I identified:
- Figure S1.2, line 27. “Increased” not “increases”
- Line 444: “dating to” not “dating of”.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Summary:
In this paper the authors investigate the fine-scale population structure of northwestern France using a combination of haplotype sharing methods (e.g. chromopainter and finestructure, and IBD analysis) and rare variant analysis incorporating both modern and ancient individuals. Their analysis reveals extensive subtle population structure in northern France, comprising 154 genetic clusters, which segregate geographically and at coarser levels (e.g. k=3 and k=18 for TVD tree) show relationships with the distribution of surnames and linguistic groupings across France. The authors posit that their observations support the hypothesis that language and rivers have played some role in shaping the genetic structure seen in northwestern France. In addition the authors explore changes in demography across time and ancestral contributions from Europe. Notably they discover putative evidence of a minor population contraction in Brittany aligning to the period of the Black death, and a strong ancestral link between Brittany and Ireland and West-Britain. Finally the study explores the relationship between modern and ancient individuals. They find striking evidence of genetic continuity in France from the Mediaeval period to present day and observe substantial steppe-related ancestry north of the Loire, with particularly high allele sharing between people from Western Brittany and individuals associated with the Bell Beaker culture.
I believe the reported analysis will be of interest to readers of Nature Communications as it demonstrates several important findings about the population history of northwestern France. The manuscript is well written and the majority of analyses appear robust, however I have a number of questions and suggestions that I would like to be addressed prior to publication.
Signed:
Ross Patrick Byrne

We thank the reviewer for his positive general assessment about the impact and quality of our manuscript. Comments and suggestions are addressed below one by one.

Major comments:
- Correlation between clusters and dialectal areas (Lines 241-254 and S figures S1.11-S.1.12: I am unclear how you are relating your clusters to changes in dialectal areas in this section. As far as I can see there is nothing in the methods section describing what metric is used to correlate genetic clusters with dialectal areas, and the supplementary figures are left unexplained (no figure legends, unclear what exactly these are showing). To itemise my questions:
1.) Are you quantifying overlap between clusters and dialectal regions, or have you made a qualitative observation of overlap? (if the latter I think this should be stated in the text as it is currently presented as a result rather than an observation).

The comparison between genetic structure and dialectal areas is mostly qualitative. We have first found positive correlation between surname-based distances and pairwise FST values across North-western France (revised Fig. S11). We then explored visually the distribution of genetics and linguistic features (revised Fig. S12 & S13) as well as other possible cultural features and discuss it for a larger audience. At this fine geographical level, a quantitative model is deemed very difficult because single dialectal features do not necessarily define a "linguistic area". We have changed the main text:
"In an attempt to shed light on the origin of the observed genetic structure, we used partial Mantel tests to correlate pairwise F_{ST} values and surname-based distances (Fig. S11). We found such correlation to be high even after correcting for geographical distances (Partial Spearman correlation = 0.68). To illustrate this finding, the "Leon" and "Malo-Rennais" cluster locations match those of the surname clusters Brest/Morlaix and Saint-Malo/Dinan, respectively (Fig. 1c). Concordant with an

association between genetics and language, we observed that the cluster "Bretagne-Centre" overlaps with the dialectal area featured by the usage of two initial consonants - the aspirated [h] instead of an unaspirated, and the alveolar fricative [z] instead of [s] (Fig. S12 and Supplementary Discussion for more details). Similarly, the cluster "Cornouaille" apparently overlaps with the area with palatalization of -h- [h] into -y- [j] (Fig. S13; see Supplementary Discussion for more details). Although these observations are not supported by statistical tests, they are consistent with a spatial correspondence between genetic clusters and dialectal areas."

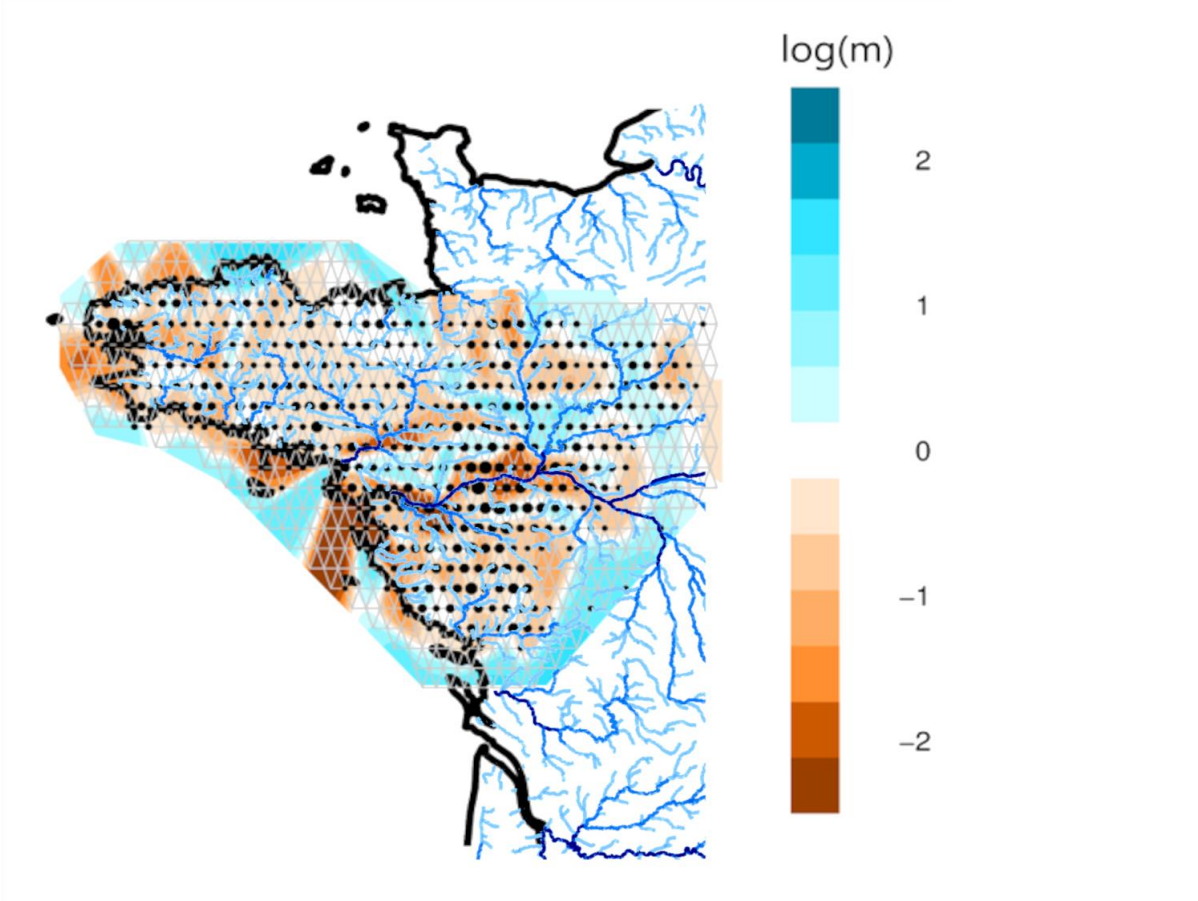
2.) What are the maps in S Figures 1.11 and 1.12 actually showing? Are these just maps of the dialectal regions or are they showing some measure of correlation with your clusters (The phrasing "in correlation with cluster [x]" in the figure titles suggests to me you have somehow measured correlation between dialectal area and genetic cluster, but the figures appear to be just regional maps outlining the dialectal areas).
If there is no formal quantification of the correlation between the dialectal regions and your clusters I would suggest doing the following to allow the reader to assess the degree of correlation for themselves:
i.) Plot all the samples from just the relevant cluster (e.g. Bretagne-Centre for S1.11 and Cornouaille for S1.12) on the supplementary dialectal maps to show how well they overlap with the dialectal areas.
ii.) Report the percentage of samples in the relevant clusters falling within the main boundaries of the dialectal areas. (You could define a polygon covering the majority of the dialectal area and perform a point in polygon analysis to count the number of samples with geocodes falling in that region).
Another suggestion would be to formally test the relationship between the geographic region corresponding to the dialects and your genetic clusters, following the framework laid out in Supplementary note 9 in Bycroft et al. 2019 (https://www.nature.com/articles/s41467-018-08272-w#Sec6)

We apologize for the lack of clarity of the former Figures S1.11 and S1.12.
As explained above, these two figures report visual explorations of similarities between the spatial distributions of genetic clusters and linguistic features at fine scale. We have now modified and expanded the figure legends as well as the related section in the main text (*Fine-scale genetic structure, linguistics and geography*).
As suggested by the reviewer, we have also added the FineSTRUCTURE results from the relevant clusters next to the dialectal maps in the new Figures S12 and S13, in order to help the reader in inspecting their overlap with the dialectal areas.
We agree that it would be interesting to formally quantify the spatial correlation between genetic and "cultural" (in a broad sense) information, but these rather complex analyses at very fine scale would require dedicated efforts beyond the scope of the paper.

- Quantifying the effects of rivers on structure: The correlation between cluster borders at finer scales and rivers other than the Loire is interesting, but I wonder if the notion that rivers influence structure can be quantified a little more solidly than visual inspection of these boundaries:
i.) Could you test if neighbouring clusters on either side of a river typically show stronger Fst than neighbouring clusters not separated by a river (at a similar geographic distance)? (as you observed for the Loire at k=3 when considering Fst for EBP vs SLO compared to EBP vs WBR).
ii.) Alternatively could you fit a geo-genetic model such as EEMS (Petkova et al. 2016) or MAPS (Al-Asadi et al. 2019) and examine if there are lower estimates for migration across these rivers? The idea here is to determine if the differentiation between samples across a river is greater than expected by a pure isolation-by-distance model.

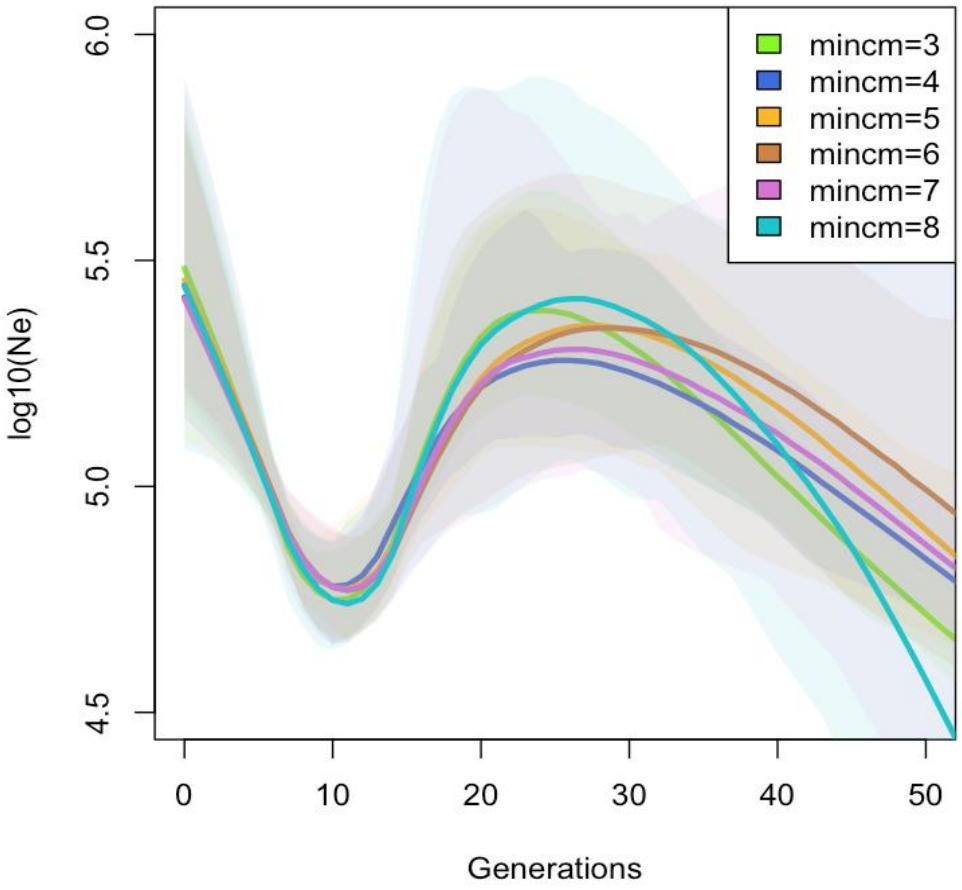
Either of these analyses would lend more weight to the statement that rivers other than the Loire influence structure.

Testing the Fst among these three large groups is relatively straightforward, as there are clearly two clusters on one side and another on the opposite side of the Loire River. However, the geographical structure doesn't readily lend itself to a similar test for smaller clusters. In line with the second suggestion, we utilised EEMS and observed variable effective migration rates across Western France, supporting the hypothesis of a departure from Isolation by Distance.
It's important to note that differences in allele frequencies between clusters (thus Fsts) may not be easily detected by fitting a geo-genetic model. Nevertheless, we observed that the main features are generally consistent.
In our EEMS analysis, we noted a significant decrease in migration across the Loire River (dark blue). Please note that rivers are coloured according to the long-term average discharge. Some migration "barriers" coincide with relatively large rivers such as Aulne (separating Cornouaille and Léon) and Laita (between Cornouaille and Bretagne-Centre), primarily in the westernmost part of Brittany. The overlap between low migration rates and rivers is not systematic though. For instance, while the Vilaine River coincides with an EEMS predicted barrier, its downstream part is in the centre of one cluster (Guérande, close to the sea). However, it matches the North-West border of cluster “Nantes” were migration rates were inferred to be lower than expected.
We are aware that we are here reporting empirical observations that will need more in-depth analysis. We feel however that this more systematic and formal analysis is beyond the scope of the present paper. The discussion was therefore modified in order to make it clear that we report observations and propose hypotheses that will require further investigations.



- Further analysis of IBDNe bottleneck in EBP and SLO: Based on your simulations of different demographic scenarios you suggest that population structure could generate a similar IBDNe curve dip as you observe in EBP and SLO without a real bottleneck. Seeing as you have identified structure in these regions (clusters at finer finestructure and TVD-tree splits), it would be worthwhile to rerun IBDNe on the largest subcluster from these regions, which should be less structured (i.e. take the biggest subcluster within EBP at k=39 or k=154; assuming this is large enough to be run in IBDNe) and see if the trend remains? If it does this would give more confidence that the bottleneck is real (Although I note that your gene-flow simulation also produced a similar curve).

We thank the author for the suggestion. We ran IBDNe on a subset of individuals of the EBP cluster, i.e one of the clusters at k=154 (n=196 individuals), and we observed the dip (see the figure below) in the Ne profiles regardless of the IBD segment length used for the Ne inference. Indeed, this allows us to "almost" exclude the effect of population structure on the Ne trajectories for the EBP cluster. Nevertheless, the coancestry matrix indicates that the *Maine Anjou* cluster (fig S1.7) has received recent ancestry from most of the neighbouring clusters and therefore, we cannot totally exclude the impact of recent migration on the coalescent rates as suggested by our simulation study and pointed out by the reviewer. To be more precise in the text the corresponding lines (lines 286:289) can be read now as:
"In addition, our computer simulations suggest that alternative demographic scenarios involving recent migration or population structure within the EBP and SLO clusters generate similar Ne profiles (see Supplementary Discussion for details, Fig. S27). Given that a large part of the "Maine-Anjou" cluster has received genetic ancestry from most of the surrounding areas (Fig. S7, coancestry matrix) and the region south of the river Loire encompasses the largest number of subclusters, such alternative scenarios cannot be excluded.".



- Globetrotter analysis results incompletely reported: You state in the text that "all seven French populations showed evidence of admixture (P<0.0001)" in your GLOBETROTTER analysis, and describe in the methods section the process by which you infer confidence intervals for admixture dates etc, however these dates are not reported anywhere (unless I have missed something). Additionally I believe your p-value is actually P<0.01 (see ii. below):
i.) Could you provide some sort of a table (supplementary would be fine) summarising the key output of GLOBETROTTER for each population tested (e.g. "best-guess" conclusion for admixture, Major and minor admixture sources, confidence interval for event timing (generations or years)). The variation in dates and number of admixing sources may be of interest to readers.
ii.) I am unclear how you get P<0.0001 with only 100 bootstraps (100 bootstraps reported in your methods section). To quote the GLOBETROTTER manual "the proportion of inferred date(s) that are <= 1 or >=400 give you the p-value for any evidence of detectable admixture." This means that with 100 bootstraps the lowest p value estimate you can get is P<0.01 (i.e. less than 1 in 100). To report a p<0.0001 you would need 10000 bootstraps…

We thank the reviewer for the suggestions. The premise of using GLOBETROTTER ancestry profiles was to confirm whether haplotype and allele-frequency methods provide us with similar ancestry landscapes. Given that they generally do, we have decided to focus more on the supervised admixture analysis. Nevertheless, we reported the GLOBETROTTER results as they can be directly compared with ancestry profiles previously reported for other European countries. That said, we have added the Table S1 in the SOM (see below) with the details of the estimation of the admixture dates and confidence intervals, putative sources of admixture and models of admixture.

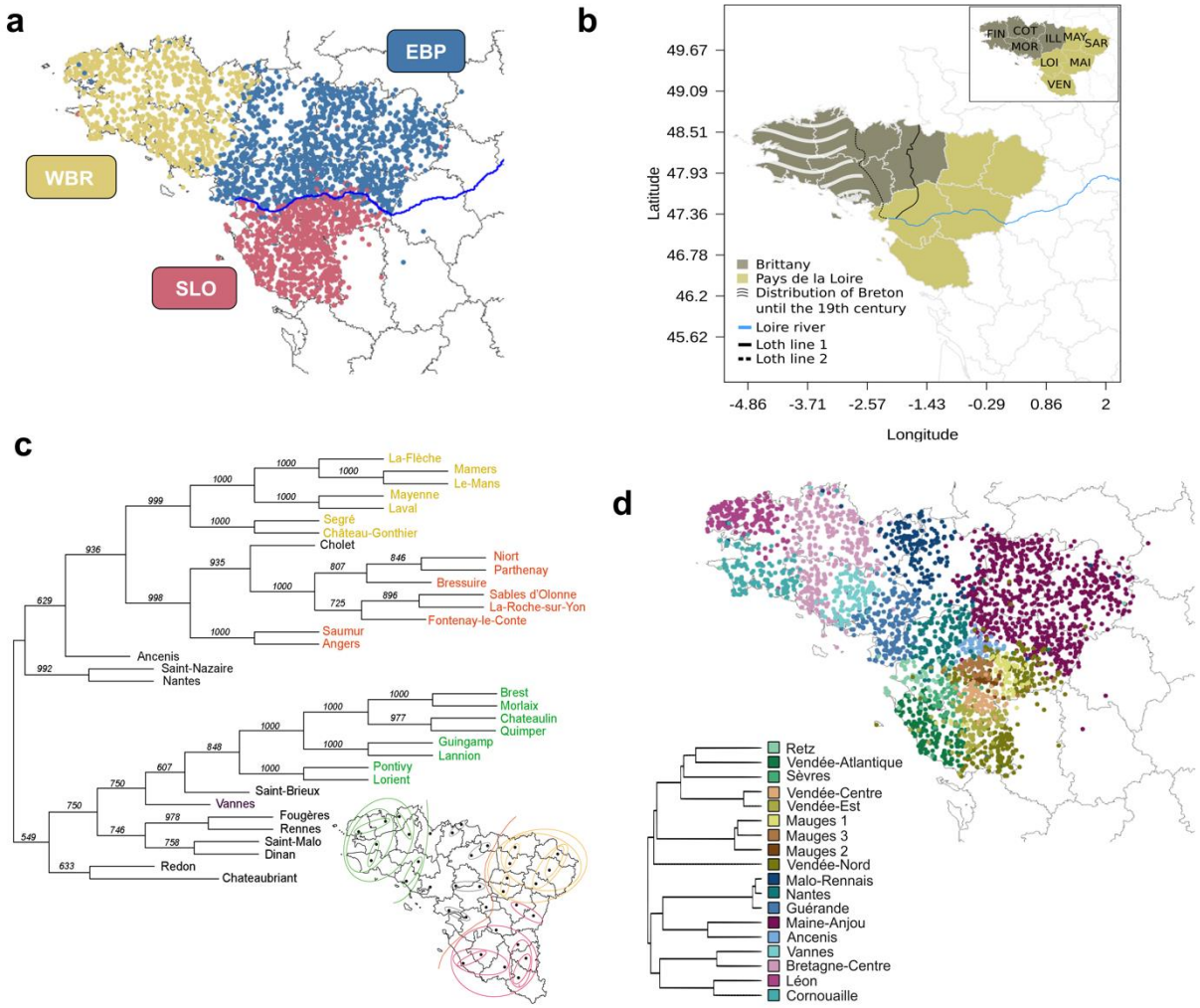
				Admixture event 1			**Admixture event 2**		
French region	**Best Fitting Admixture Model**	**Goodness of fit (R^2)**	**Time of admixture in ya [95% CI]**	**Minor Component**	**Major Component**	**Ratio**	**Minor Component**	**Major Component**	**Ratio**
Brittany (BRI)	One-date-multiway	0.62	1047 [858-1244]	Italy	Kent	0.32:0.68	Kent	Kent	0.43:0.57
Pays-de-la-Loire (PAY)	One-date-multiway	0.70	1337 [1163-1554]	Italy	Kent	0.49:0.51	Kent	Spain	0.48:0.52
Centre et Val-Loire (CEN)	Unclear signal	0.19	NA	NA	NA	NA	NA	NA	NA
Normandy (NOR)	One-date-multiway	0.30	1035 [733-1293]	Denmark (Denm)	Italy (Ital)	0.47:0.53	Ireland (Irel)	Norfolk (Norf)	0.24:0.76
Hauts-de-France (HAU)	One-date	0.65	1105[951-1354]	Italy	Kent	0.49:0.51	NA	NA	NA
Grand Est (GRA)	One-date	0.50	1148 [925-1438]	Kent (Kent)	Italy (Ital)	0.47:0.53	NA	NA	NA
Nouvelle-Aquitaine (NOU)	One-date-multiway	0.35	1540 [887-1914]	Kent	Spain	0.31:0.69	Kent	Italy	0.42:0.58

Concerning the p-values, we thank the reviewer for pointing this out. There was indeed a mistake/typo and the p-value has now been changed to < 0.01. The sentence (lines 359-361) now reads as: "We found that all of the seven French populations showed evidence of admixture ($P < 0.01$, computed from the proportion of inferred dates outside of the limits, i.e., <=1 or >= 400)."

Minor comments:

-Figure 1 and lines 172-180: You define the WBR, EBP and SLO clusters clearly in the text, but it may help readers to interpret the figure and results more readily if these clusters are also labelled in figure 1a (Simplest option is including a key in panel a mapping WBR to yellow, EBP to blue and SLO to red).

We thank the reviewer for the suggestion. We have now added the labels on Fig. 1a for clarity, and have changed the legend accordingly, adding the sentence : WBR stands for Western Brittany, EBP for Eastern Brittany and *Pays-de-la-Loire* and SLO for South Loire.



- What is meant by "levels of clustering" Line 165-167: I am unclear what is meant by the statement "At this finest scale, levels of clustering were similar to those previously found for Spain (3), and larger than those previously reported for France (31) and Great Britain (5)." Do you mean the number of clusters was larger or the size of the clusters? If you mean the number of clusters could you comment on what your interpretation of this is and why it is important to report? I would expect that the raw number of clusters found by finestructure is somewhat a function of the sample size, and not necessarily a good measure of the degree of structure in the population being analysed. For example, the previous study of France cited in this comparison uses samples from all regions of France, but has a much smaller sample size which likely explains the lower number of clusters (If sample sizes were equal we would probably expect to see more structure in the whole country than just the northwestern region).

We thank the reviewer for pointing this out. Indeed, the number of clusters identified by fineStructure heavily depends on the number of samples analysed (Lawson et al 2012). One way to go around it would be to explicitly compare the ratio of clusters/sample size. In such a case, we would assume that there is a linear association between sample sizes and clusters, which is not necessarily true. Therefore, we decided to remove the sentence "At this finest scale, … and Great Britain (5).", from the main text.

-IBDNe simulations: In your simulation study you state "We ran the IBDNe software on our data only using IBD segments that were 4cM in size." I think you mean "greater than 4cM" here

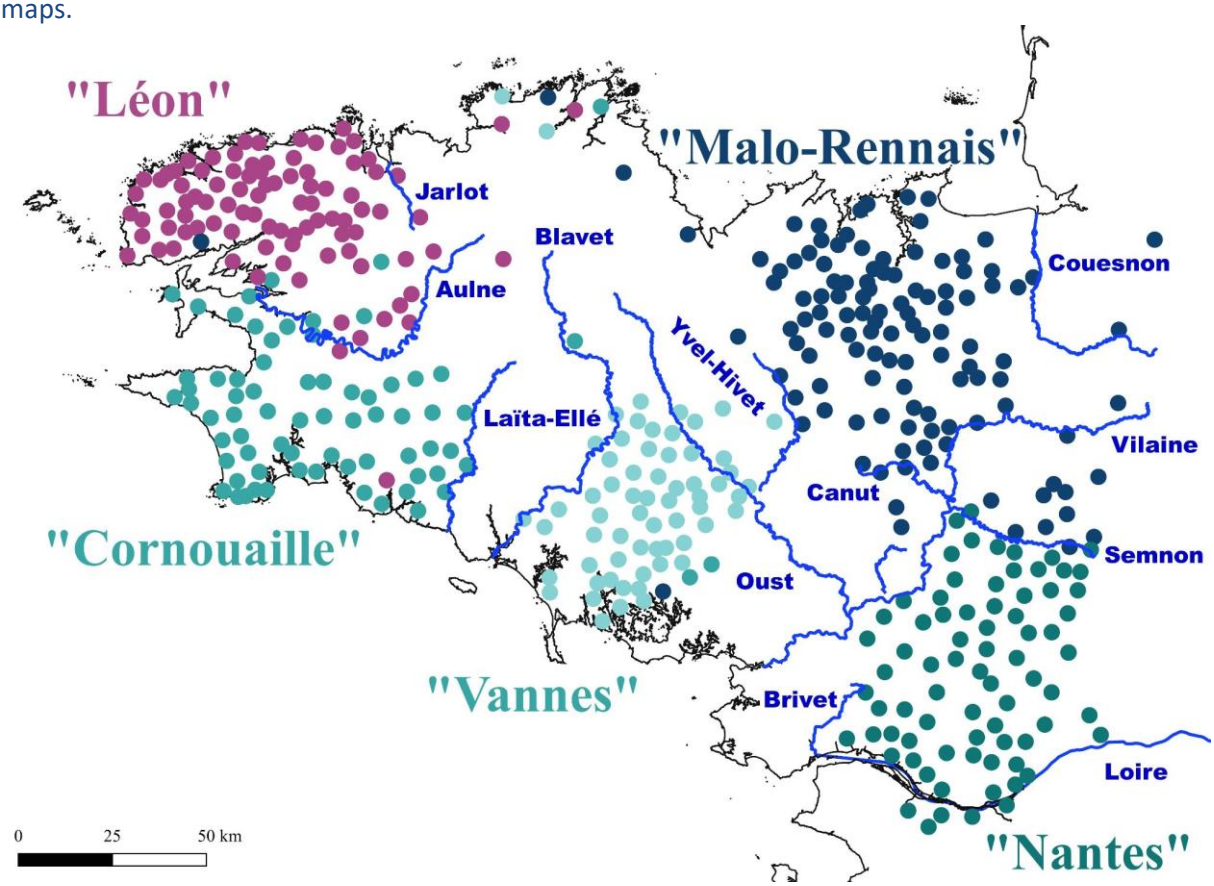
Thanks for pointing out this. We have corrected the supplementary text accordingly. It now reads: "...IBDNe software on our data using only IBD segments greater than 4cM."

-Line 286 in main text referring to fig S1.15: I think you mean "minimum centimorgan length" here not "chromosome length".

Thanks for spotting this. We have corrected the main text accordingly (line 286): **"**observed across different thresholds for the minimum centimorgan (cM) length (Fig S15b-d)**".**

-Supplementary figure 1.14: This figure could be made clearer by increasing the width of the plot lines in the map for the borders and rivers. At present the thin lines make it hard to distinguish borders from rivers. In addition the colour scheme for clusters in panel three is quite hard to distinguish due to the use of similar shades of green being used. You might consider changing the colour scheme or alternating plot characters (i.e. circles, triangles etc) to make this figure clearer.

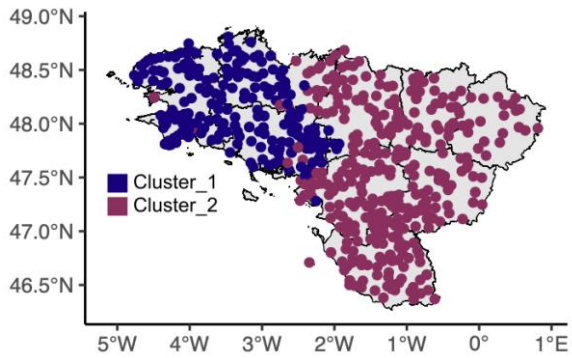
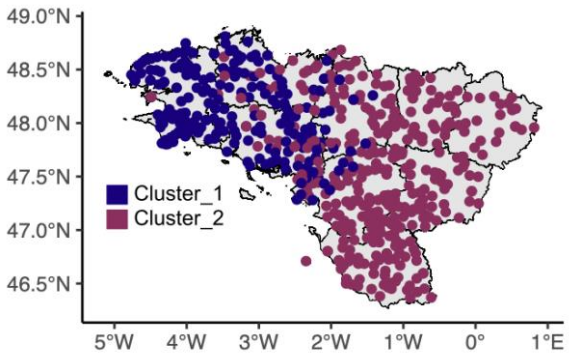
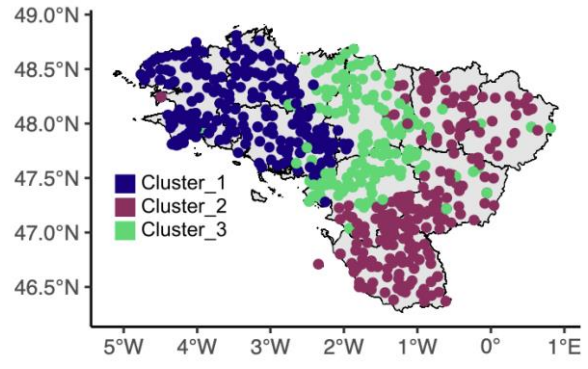
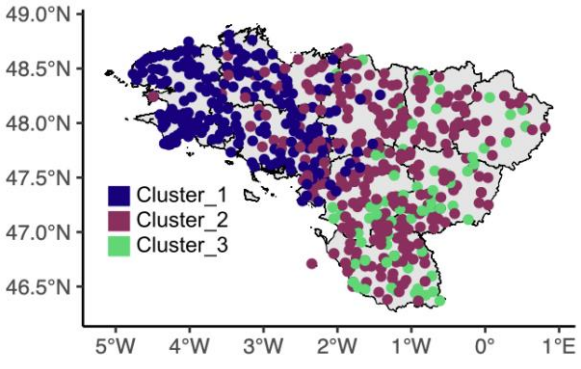
The Figure has been improved according to reviewers' suggestion. We show here one of the three maps.

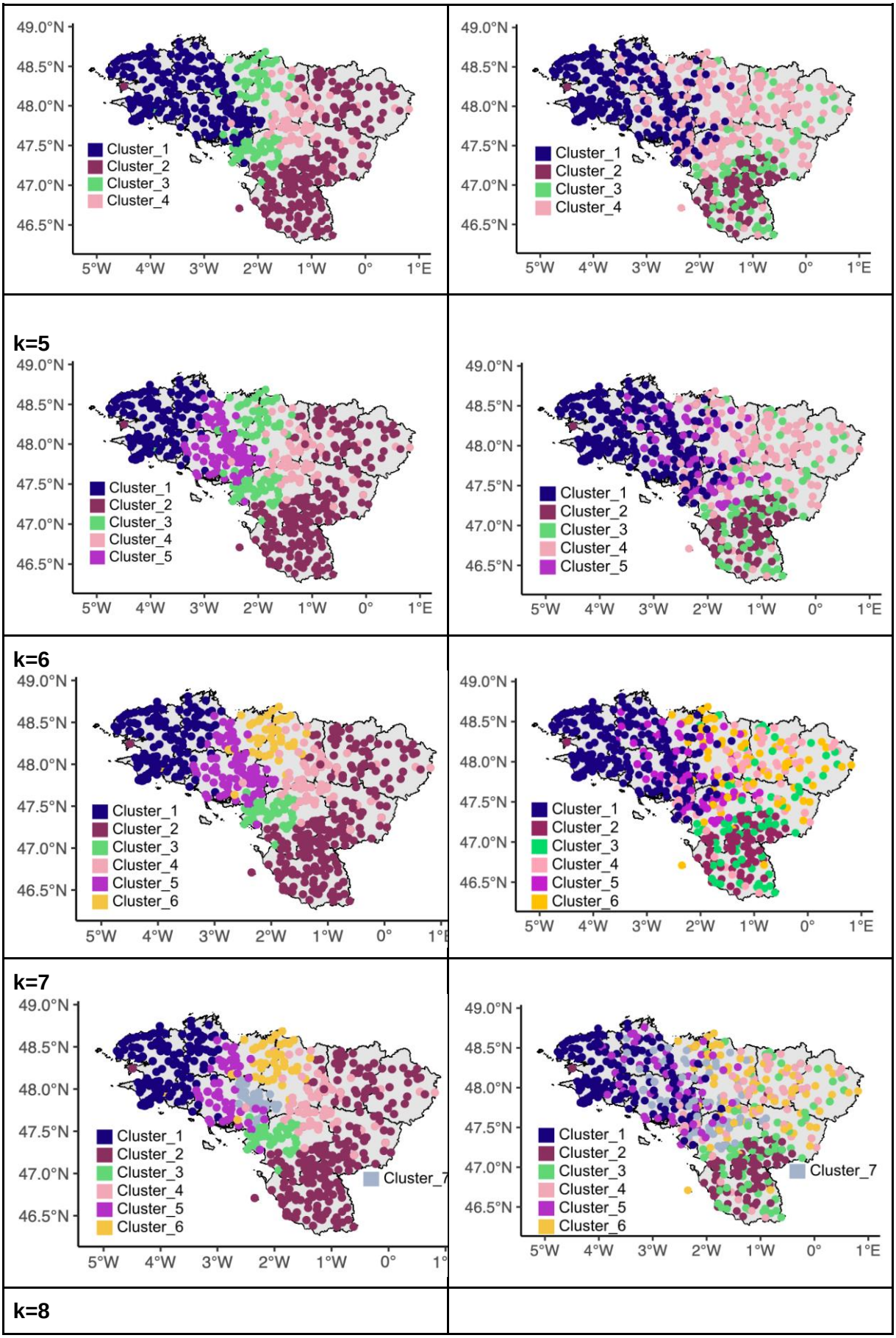


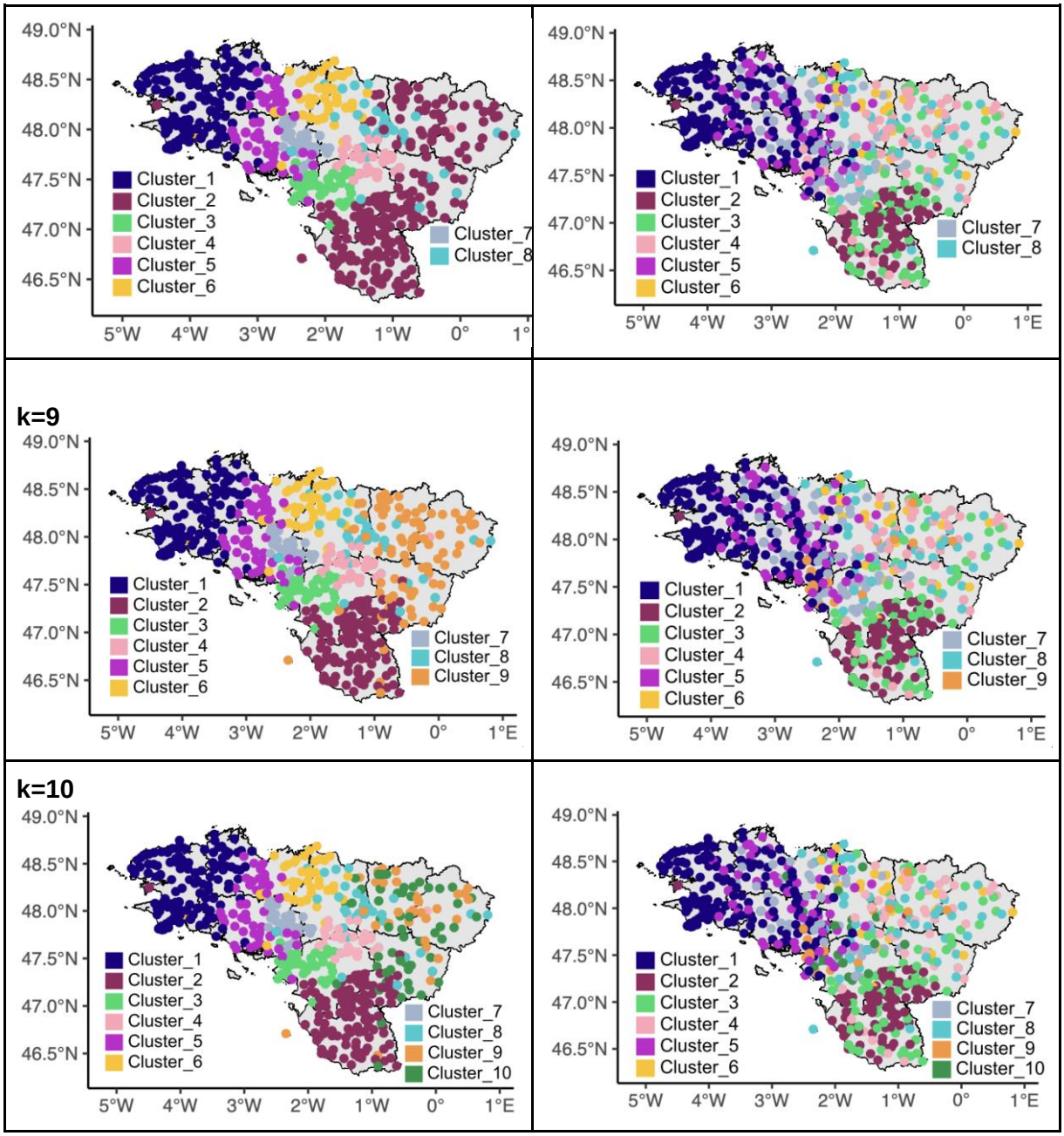
- Plots of finescale structure based on rare variants: In figure 2 have plotted your rare variant sharing clusters using picharts per département rather than plotting each individual on the map coloured by their cluster (as in finestructure map plots like figure 1 and S.1.1 etc). Is there a reason for this visualisation choice? I'd be curious to see the raw geographic distributions of the clusters but understand that this might be a neater visualisation given the likely overlapping cluster boundaries.

The main reason for using pie charts is related to the (5-times) lower sample size in the WGS dataset in comparison with the SNP array. Because of that and mainly for higher k values, the points are spread out and mixed up, hindering visualisation, as the reviewer predicted. We decided to keep the pie chats on the main text but we added an extra figure on the SOM with those maps (now Figure S18) as for low k values they are still informative. For instance, it is easier to understand that clusters based on doubletons are far more geographically constrained than for higher mac values than variants with an allele count >3.

The new figure looks like this:

MAC 2	MAC 3-10
k=2 	
k=3 	
k=4	


k=5
k=6
k=7
k=8
49.0°N
48.5°N
48.0°N
47.5°N
47.0°N
46.5°N
5°W
4°W
3°W
2°W
1°W
0°
1°E
Cluster_1
Cluster_2
Cluster_3
Cluster_4
Cluster_5
Cluster_6
Cluster_7



The figure has been added to the main text that is as it follows:

"Interestingly, we identified a genetic component restricted to the *départements* located south to the river Loire for *k*=4 (cluster_2) for MAC 3-10 (Fig. 2, Fig. S2.1 and Fig. S18). For MAC 2 alleles we found this cluster only at *k*=9 (cluster_2). Assuming that alleles with MAC 2 (minor allele frequency ≈ 0.0016) tend to be more recent, these results suggest that population structure does not result from reduced gene flow between the northern and southern shores of the river in the very near past (average doubleton age ~500 years (39)). Consistently, we found no significant differences in surname distributions between the riversides across *arrondissements* crossing the Loire (data not shown). Clustering patterns within Brittany are, on the other hand, consistent across the full range of allele counts, indicating that population differentiation associated with traditionally Breton-speaking groups has persisted to modern times.

In general, with MAC 2 alleles, increasing *k* from 3 to 10 assign individuals from neighbouring *départements* into 7 additional geographically restricted clusters (Fig. S2.1 and Fig. S18), suggesting similar patterns of population structure as found with fineSTRUCTURE (Fig. 1d). Although an exhaustive comparison with fineSTRUCTURE results is beyond the scope of this study, this general concordance in

clustering patterns emphasises the power of rare variants to infer fine-scale population structure. With MAC 3-10 alleles, increasing *k* tends to generate smaller clusters with relatively large geographical distribution, likely reflecting a relative lack of resolution to detect very recent population structure (Fig. S17 and Fig. S18).*"*

Reviewer #2 (Remarks to the Author):

The manuscript titled "Genetic population structure across Brittany and the downstream Loire basin provides new insights on the demographic history of Western Europe" presents a comprehensive investigation into the genetic history of France. To do so, Alves, Giemza et al. generated novel genomic data for approximately 4,000 individuals from the northern half of France and analyzed it, together with publicly available data for present-day and ancient European individuals, with state-of-the-art methods in population genetics. The authors skillfully employ a detailed population genetics analysis and integrate it with linguistic data, geographical data, family names, and historical context to unravel the intricate population dynamics that have shaped the present-day genetic landscape of Northwestern France and describe the population structure of the region in detail.
The manuscript's strengths are notable right from its outset. The authors identify the critical role of Northwestern France as a geographical meeting point of three main ancestral European populations, namely Western Hunter-Gatherers, Early European Farmers, and the Yamnaya Eurasian Herders from the Steppe. This central positioning lends significance to the study of the genetic history of the region. The comprehensive dataset of approximately 4,000 present-day individuals from Northern France, along with the inclusion of ancient DNA data from six medieval individuals from Western France, highlights the comprehensiveness of the study's sampling strategy – another notable strength of this work. This approach not only allows for the exploration of fine-scale population structure (as the authors thoroughly examine in this manuscript) but also presents an opportunity to address the impact of historical migration events (including events during the Pre-History, Antiquity, and the Migration Period). I find it particularly interesting, for instance, the authors' insights about the relationship between Brittany and Ireland based on their genetic data.
In my point of view, all methods were used to the highest standards, and all conclusions are fully supported by the data. I have not identified any major weaknesses with this manuscript; however, I have a few minor comments below and would like some clarifications about the ancient DNA (aDNA) data and analyses. Knowing the previous works of some of the authors, I believe the aDNA analyses were correctly performed. I do, however, think that more details are needed. I apologize in advance if the issues I bring up are already explained in the manuscript and I have missed them. If so, I would appreciate it if the authors could point me to the relevant sections in the text.
In sum, I believe the take-home message from this paper is exciting and of broad interest. This work presents a technical tour de force and reports a broadly useful data resource. This paper, when published, will be of great interest to those in the fields of population genetics, anthropological genetics, human genetics, and genomics. The dataset generated by this work will also be an extremely valuable resource for these fields.

Thank you for the opportunity to review this excellent manuscript.

Best,
C. Eduardo Amorim

We thank the reviewer for the very positive and constructive reviews. We addressed them one by one below.

COMMENTS ON THE aDNA ANALYSES

I have a few concerns related to the authentication of the aDNA data and whether the necessary filtering of the data was performed.

1. Were the aDNA data authenticated somehow? I see a table describing the depth of coverage and contamination rates; however, I did not find any mention as to whether post-mortem damage (PMD) patterns in the DNA sequencing reads were assessed. In my point of view, PMD patterns must be assessed in a study such as this. If the PMD patterns were assessed in the aDNA samples, are they typical of aDNA? I kindly request the authors to provide some information about this, perhaps a figure or additional columns to Table S2.1 describing the percentage of terminal bases with C-to-T and G-to-A transitions.

We thank the reviewer for bringing this in. Yes, the samples were authenticated and we have now added two new columns to the now Supplementary Table S2 "3' and 5' damage %" and "Average Read length" to provide more information about the quality of the samples. The proportions of damage range from 9.1 to 25.6% which is pretty much expected for samples of the Mediaeval time period (Rodríguez-Varela et al., 2023 Cell). The table looks like this:

Table S3. Summary of the ancient DNA Mediaeval samples from France sequenced in this study

Samples ID	Dates*	Place	% endog.	Mapped reads	3' and 5' damage %	Average read length	Genome cov.**	Sex***	X-contam. (SE)	mtDNA contam.	total sites HOA	mean cov.
fra001	375-541 cal. AD	Saint Lupien Rezé	0.199312	271,093,228	16.9-17.3	75.38	0.48	XX	NA	0.01	247020	0.551608
fra004	414-553 cal. AD	Saint Lupien Rezé	0.057764	15,241,917	14.2-15.0	72.26	0.28	XY	0.04367 (0.009)	0.02	169051	0.344645
fra008	943 - 1024 cal. AD	Chaussé Saint Pierre - Angers	0.018284	5,983,204	25.1-25.6	55.97	0.08	consistent with XY but not XX	0.22336 (0.000)	0.01	64438	0.116075
fra009	414-548 cal. AD	Chaussé Saint Pierre - Angers	0.276266	26,781,466	23.6-23.8	70.51	0.36	XY	0.03787 (0.007)	0.01	199735	0.421043
fra016	600-700 AD	Chémeré	0.480058	234,557,159	24.3-24.5	56.46	3.43	consistent with XY but not XX	0.014 (0.001)	0.01	574582	4.153501
fra017	600-700 AD	Chémeré	0.020224	6,559,927	9.1-11.1	80.88	0.14	XY	0.10887 (0.000)	0.01	92736	0.170474

* cal. AD = Radiocarbon calibrated Anno Domini (AD); AD = Anno Domini based on the archaeological context

** after MapQ 30 filtering

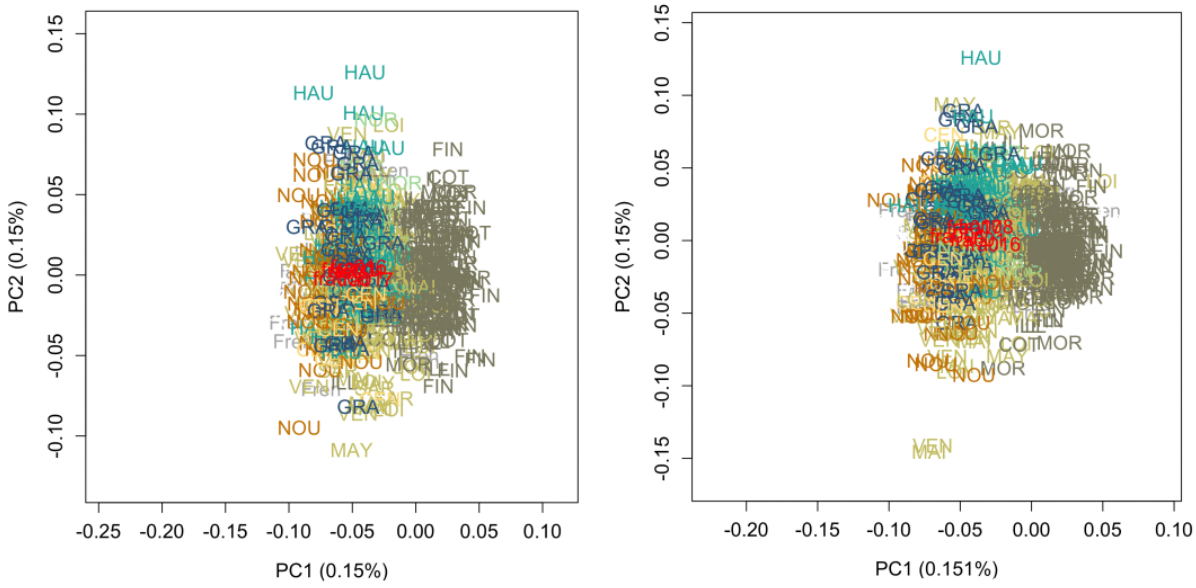
*** estimated biological sex from alignment data (1).

%endog. - % of endogenous DNA.

2. On a related note, how PMD was taken into account when calling variants? I'm assuming the samples were not UDG-treated. If so, were transitions included in the analyses? If not, why? Please consider mentioning whether UDG treatment was performed and how.

We thank the reviewer for the remark. Indeed, no UGD treatment was implemented and we used all the sites, including transitions present in the HO array. Although we are aware that analyses in population genetics should be performed only on transversions, when no UGD treatment is performed, we think that by focusing on a set of known SNPs, as those present in the HO array, minimises potential aDNA-related biases. Indeed, the premise of focusing on HOA SNPs lies in maximising the number of samples available while minimising, at least to some extent, the potential bias associated with aDNA. To explore this further we performed the same PCA as in now fig. S24 using only transitions (85,000 randomly sampled SNPs) and only transversions (85,000 randomly samples SNPs) and we do not see considerable differences when plotting the ancient samples onto the PCs built on the present-day French whole-genomes. The only thing we observe (see below) is a reduced differentiation between the different French groups likely due to the lower number of SNPs (85,000 out of 471,411 SNPs).

Transitions (85k SNPs) Transversions (85k SNPs)



3. There is no mention of the high (>10%) contamination estimates for fra008 and fra017 (X-contamination column Table S2.1). Reading the manuscript, nothing strikes me as a bias introduced by contamination, but I wonder if the authors could perhaps reassess their conclusions in light of the potential biases introduced by contamination. Perhaps just a brief mention of it in the manuscript should suffice.

Thanks for pointing this out. Indeed, we have high contamination estimates from the X-chromosome (we have now added the SE estimated with ANGSD on Table S3). Nevertheless, mtDNA-based estimates are not consistent with those estimates and we suspect that the differences just reflect the low coverage in these two samples. Globally, as the reviewer noticed, we did not find any particular sign of contamination. The sample whose PCA suggests to have affinities with present-day North Africans is not among the ones with >10% of contamination. On the other hand, we could think that contamination with present-day DNA would make the ancient samples closer from the present-day samples and erase any signal of discontinuity between Mediaeval and present-day French. Indeed, allele sharing analyses points towards a general continuity between the present and the Mediaeval period. However, samples from early Mediaeval (fra001 and fra004, contamination <10%) and samples from a later Mediaeval period (fra016 and fra017), among which we have one of the samples with

>10% contamination, also exhibit continuity, suggesting that the close relationship between Mediaeval and present-day French is real and not driven by contamination. As suggested by the reviewer we added a sentence explicitly mentioning the levels of contamination found in our samples (lines 423-426). It reads as: "Although two of our samples (fra008 and fra017) appear to have levels of contamination >10%, this is not the case for fra009 (contamination ~4%, Table S2), which led us to exclude contamination as a source of its outlier position." and line 625 "(fra016 and fra017, with the latter showing high levels of contamination ~ 10%)."

4. More details are needed in the aDNA-related methods description.

We thank the reviewer for the suggestion (including reviewer's point 5 below). We have now moved the section of the methods related to the aDNA ("aDNA Mediaeval samples" and "aDNA library preparation and bioinformatic processing") to the beginning of the Methods, after the description of the present-day genotyping and whole-genome sequencing data. We also added the details according to the reviewer's suggestions. We address each specific comment below.

a. Was DNA extraction and library prep done in a clean aDNA-dedicated lab?

Yes, they were. We added the following lines 832-833 in order to clarify this. "DNA extractions and library preparations were performed in the dedicated clean aDNA-laboratory at the Uppsala University, Sweden."

b. Please add a supplemental figure with the radiocarbon dating results.

We added whenever possible to the Supplementary Online Information - Section Supplementary Archaeological Information. See four answers below for details.

c. Were the libraries single or double-stranded?

All the libraries were double-stranded. We added this information in the methods (lines 831-832): "All DNA libraries were double-stranded. DNA extractions and library preparations were performed in the dedicated clean aDNA-laboratory at the Uppsala University, Sweden."

d. I believe paired-end reads were used, but I couldn't find any clear mention of it. (Forgive me if I missed it)

The sequencing chemistry information can be now found in lines 833:837. It reads as: "The DNA libraries were sequenced in two batches and over multiple lanes, first as a pilot run at the SciLife Sequencing Centre in Uppsala, Sweden, using Illumina HiSeq 2500 with paired-end 125 bp chemistry, and later in more depth at the CNRGH (Evry, France) using Illumina HiSeq X and with a paired-end 150 bp chemistry."

e. Were duplicates removed? The last paragraph of page 27 describes a methodology that I am unfortunately not familiar with. Consider rephrasing it to add more details.

We thank the reviewer for pointing out the lack of details. We changed the sentence describing the processing of the BAM file (lines: 868-872). It now reads: "BAM files were merged to a per sample library level using the merge command in Samtools version 1.5 (76) followed by PCR duplicates removal (reads with identical start and end positions were identified and collapsed) using a modified version of FilterUniqSAMPCons_cc.py, which ensures random assignment of bases in a 50/50 case, as described in (77)."

Table S2.1 is a bit confusing. Are these dates based on radiocarbon dating or archeological context? The manuscript mentions the former but, to the best of my knowledge, the format is unusual for radiocarbon dates. Also, I think the column "HOA sites" is distracting because, I believe, it has the same value for every row – consider excluding this column and adding the of number sites to the table description. Last, is the "mean cov." column describing the mean coverage in the HOA SNPs or else? – unclear to me.

We thank the reviewer for the remark. Concerning the radiocarbon dating, samples "Samples fra001, fra004, fra008 and fra009 were dated using radiocarbon methods and estimates vary from 375-1024 cal. AD (Table S3)." and samples: "fra016 and fra017 - are based on the archaeological context". We added the estimation plots, whenever available, in the Supplementary Online Information - Archaeological details. Most of the samples have been dated before our study (with the exception of the fra008). For the sample fra009 we extracted the information from the official reports (in French) done by the personnel of the "Institut National de Recherche archéologiques préventives". In these reports all the dates are in the calibrated years AD. We have contacted the archaeologists to see if we have the Carbon dates for the samples fra016 and fra017, but we have had no answer yet as they are currently doing fieldwork. To make this clear in the text we have changed the methods as follows: "Samples fra001, fra004, fra008 and fra009 were dated using radiocarbon methods and estimates vary from 375-1024 cal. AD (Table S3). This interval corresponds to the early and High Mediaeval periods. The dates of the other two samples - fra016 and fra017 - are based on the archaeological context (66). Out of the six ancient individuals, four (fra001, fra004, fra016, fra017) were sampled south of the river Loire while two were sampled north of the Loire (fra008, fra009, Fig. 4)." We have also changed the main text as follows: "to increase our resolution in detecting changes in ancestry during the Mediaeval Period, we sequenced DNA libraries of four ancient samples with radiocarbon dates ranging from the 375-1024 cal. AD, from *Pays-de-la-Loire* (Fig. 4a) and two samples from archaeological sites dated to the 6-7th century".

5. One small editing suggestion I would like to make is that you move the aDNA methods part to somewhere closer to the beginning of the Methods section. Although I understand this is a matter of writing style and the authors may disagree with this, I believe this would make the text more logical to the reader. For instance, you say somewhere that you use "inbred:YES" for the analysis and that implies that the data is in the pseudo-haploid format. However, as far as I can recall, this is not mentioned anywhere before that point. It is mentioned later, on pages 27 and 28.

We have moved the sections related to the aDNA samples and processing to the beginning of the methods, as explained above. Hope this helps the reading of the manuscript.

6. Last, but not least, please add the source of the ancient and present-day data you used. In my opinion, it is unfair to the authors of the papers from which you used data to not have their papers cited. Citing the AADR preprint or website is not enough (by the way I believe you should cite the preprint, not the website). I quote the preprint by Mallick et al. on bioRxiv here:
"Researchers who use the curated dataset from the AADR as the basis for analyses should cite this paper and the version of the AADR they downloaded, including a reference to that version's doi. Citing the AADR paper is not a substitute for citing the original publications that produced data, which should be specifically referenced in each publication."
I understand that Nature Communications might have a limitation on the number of citations that can be included in the main text. If that's the case, to avoid going over this number, I suggest authors cite these papers in the supplemental material, therefore acknowledging the hard work of several people who contributed the excellent and very useful resource that AADR is.

We acknowledge that it is indeed important to reference the original publications, and have added the list of publications to the Supplementary Information - Section Supplementary Discussion. We also added references to this section on the main text (for instance Fig. 4 "ancient merged dataset")

MINOR COMMENTS:

[Table 1] What does "published" in the name of the samples on this table and elsewhere mean?

We thank the reviewer for the comment. This relates to the original meta information of the HOA dataset (Reich's Lab, vs.42.2 March 2020 release). As it has no meaning in our study, the "published" was removed.

[Fig 4a] What are the numbers in the parenthesis after the name of the groups? Sample sizes, citations, or else?

We thank the reviewer for pointing out the lack of information. It represents the sample sizes. This information is now added on the figure legend. Fig .4a "The numbers within brackets represent the sample sizes."

[Page 20] The description of the SNP array genotype data generation is very detailed and mentions the exclusion of related individuals and the filtering of SNPs. The next section describes the generation of WGS data. Was relatedness also assessed within the WGS data? Were SNPs with minor allele frequency (MAF) <5% also removed from the WGS dataset? Have authors tested if there were batch effects introduced by different sequencing strategies (SNP capture vs. WGS)? Are these 856 WGS distributed evenly across the studied region? I apologize in advance if I missed this information.

We understand the confusion about the sample sizes and MAF filters and we thank the reviewer for pointing this out. Relatedness was also assessed within the WGS data, where we also excluded three samples based on the PI_HAT statistics. Then sample sizes and MAF filters were specific to each of the analyses depending whether we used the "merged-modern dataset", the rare variants or the "merged ancient dataset". We agree that batch effects can be an issue mostly in the case of rare variant analysis. Indeed, we tried to compute rare variant sharing between our WGS dataset and the 1000 GP but due to batch effects we did not report any results (data not shown). This is the reason why the rare variant analysis reported in Figure 2 was performed exclusively among the 620 genomes of individuals from North-western France from our WGS dataset. All the other analyses including the WGS are based on the SNPs present in the different SNP arrays (Human670-QuadCustom, Human1-2M-DuoCustom or Human Origins Array) used for merging with modern or a mix of modern and ancient samples. Although by merging with SNP arrays we expect to limit potential batch effects, they could still occur. However, when inspecting the distribution of the French WGS in the PCA on Figure 3a, we observe that they show larger variation than most of the other non-French samples within the SNP array. If batch effects due the different genotyping technologies were to be present one would expect to see a much narrower distribution of the French WGS. Also, if that would be the case in Figure 3c we would expect to see a larger sharing between pairs of French samples (WG sequenced) than between French samples and non-French samples. This is not what we observe.
To make things clearer we added a new figure (Figure S16 showing the distribution of all the 843 samples, and the following text (lines 800:803) in the section "*Whole-genome sequencing and variant calling*": "Related samples were identified using PI_HAT statistics computed with PLINK (vs.1.9) and excluded when values were >0.10. Individuals whose grand-parents were not born on the same *département* were also excluded leaving us with 843 samples. The sizes of the samples used in each analysis are specified hereafter and in every corresponding figure."

[LD pruning] Was the dataset LD-pruned for fineSTRUCTURE (as it was done for the PCA)? If not, do authors believe it could introduce bias in their results? In my experience, fineSTRUCTURE is sensitive to how the SNP set is filtered. I would like to know whether the authors, who are more experienced than I in this type of analysis, think this might have introduced any bias that may have affected their observations. Thank you!

The dataset was not LD pruned for fineSTRUCTURE analysis, contrary to PCA analysis. Because Chromopainter is using phased data (haplotype) which is more informative than genotype data, we kept all the "good quality" SNPs (see the filtering procedure). In our experience, keeping all the information, and especially the SNPs in LD, helps better assessing the haplotypes and hence to derive population information.

[Page 21] It is unclear why authors highlight the Viking dataset from Margaryan et al. 2020 within the data included in the AADR. Is this because this data was not included in version 42.4 of AADR? The dataset is included in AADR currently – therefore my confusion.

The genotypes for the HOA SNPs in the Viking dataset were obtained through collaboration before the data was publicly available. Meanwhile they became available. We added this information in the "*Data Availability*": The genotypes of the Viking samples for the HOA SNPs are now available in the version 54.1 of the ADDR (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data).

[Page 22] Please provide more information about the analyses performed in R. Was a publicly available R library used or an in-house script? Please cite the former. Consider making the latter available for reproducibility and transparency, if applicable.

We added in every figure legend the software package and the respective libraries used to perform the plots.

[Page 24 – RoH] This section might need to be more detailed and include the detailed parameters used for estimating RoH and IBD unless RefinedIBD has a standard setting of parameters. If so, please mention the default parameters that were used. I am not familiar with RefinedIBD but I have used other methods to run these analyses and observed that results may be sensitive to the choice of parameters such as missing data allowed, window size, number of heterozygote sites allowed within an RoH fragment, and minimum length of the fragment, among others. Therefore I believe you need to be more specific about what parameters you used here. Sorry if I missed this information from the text.

We used the default parameters of the version from 23rd December 2017 of RefinedIBD. The default values are :
the cM length of the sliding marker window (default: window=40.0)
the minimum LOD score for reported HBD segments (default: lod=3.0)
the minimum cM length for reported HBD segments (default: length=1.5)
the cM trimmed from the end of a shared haplotype when calculating the HBD LOD score (default: trim=0.15)
scale=0, ie= max{2, √[sample size]/100}

We have added to the methods sections (line 1004) "version from 23rd December 2017, with default settings". We agree that this analysis maybe different according to the method and the parameters. In order to check the robustness of our results, we carried out the same analysis, consisting in estimating the mean Homozyguous by Descent tracts length in individuals (Model Refined IBD) for the software plink and hap-ibd, proposing two sets of parameter values. For plink, the model (Model Plink Def) is

using the default parameters and a model fragments (Plink M1 --homozyg-kb 500, --homozyg-snp 50, --homozyg-window-het 2, --homozyg-window-missing 10). We chose plink as the algorithm tends to identify in general longer fragments and is more likely to capture different chunks.
We observe correlation higher than 90 % (at individual level) between all estimates of individual HBD length, which reassures us about the robustness of the analysis.

	Plink Def	Plink M1	RefinedIBD
Plink Def	1.0	0.98	0.90
Plink M1	0.98	1.00	0.91
RefinedIBD	0.90	0.91	1.00

Whatever the method, we observe increased mean ROH length when advancing westwards.

[Lines 953 – 955] This sentence seems to be incomplete.

We have now modified the sentence:
"Given sample size differences (see section Publicly available datasets of western Europeans), we down-sampled the source populations from Germany and Ireland to 350 individuals."

[Data availability] Are authors planning to make their data available? If so, how? Please clarify.

Yes, we plan to make our data available. We have added a paragraph entitled 'Data Availability':
"The WGS data on ancient DNA (bam files) are available at the European Nucleotide Archive (ENA), under the accession number PRJEB71835. Genetic data on contemporary human individuals are subjected to the French regulation on data privacy related to identifiable personal information. WGS data from the FranceGenRef panel will thus be submitted to the French Centralized Data Centre of the France Medicine Genomic Plan that is under construction. Enquiries for the use of this data can be addressed to GENMED LABEX (http://www.genmed.fr/index.php/en/contact). The modern genotypes data (PREGO data) have been uploaded to the European Genome-Phenome archive (EGA) website (https://ega-archive.org), under the accession number EGASXXXXXXXXXX. The annotation file describing the array is available at https://www.thermofisher.com/fr/fr/home/life-science/microarray-analysis/microarray-data-analysis/genechip-array-annotation-files.html (Axiom_PMRA Annotations, CSV format, Release 35). Further information about EGA can be found on "The European Genome-phenome Archive of human data consented for biomedical research"(https://doi.org/10.1093/nar/gkab1059)."

[PCA Fig S4.1] Why did authors use the HOA instead of the 1240K for this and other analyses?

The HOA included also a set of present-day human populations we were interested in for contextualising with the European landscape.

[Typos] Here are two potential typos I identified:
- Figure S1.2, line 27. "Increased" not "increases"
- Line 444: "dating to" not "dating of".

Thanks for pointing them out, they have been corrected.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

In my opinion the authors have done a great job in responding to my comments, both in their careful rebuttal and in the high quality edits they have made to the manuscript. In particular I believe the edits to the figures supporting the relationship between dialectal areas and population structure add to its clarity and message. My concerns have been largely addressed and I believe the revised manuscript will be of great interest to readers of Nature Communications.

I have only one minor comment, which I leave to the authors and editors decision:
I think that the EEMS plot provided in the rebuttal to support the role of rivers in influencing population structure in France might be worth including in the supplementary materials for completeness. However I understand that fully discussing this analysis may be outside the scope of the manuscript.

I highly recommend accepting this article for publication.

Signed:
Ross Patrick Byrne

Reviewer #2 (Remarks to the Author):

This version is an improvement of an already excellent manuscript. The revised version reflects a thorough consideration of my previous feedback. I appreciate the attention given to my concerns and the comprehensive response provided by the authors.

I am happy with the level of detail of the analyses included in this second version of the manuscript. I believe the manuscript is now ready for publication in its current form and I have no further concerns to note.

Below, I highlight one potential typo and make one small suggestion. I would appreciate it if the authors could double-check these two potential issues, but I see no need to send the manuscript for another round of review as those two comments are only minor *suggestions*.

(1) Potential Typo? Page 11, lines 408-410 – While I couldn't pinpoint the exact issue, it seems there might be a missing word or parenthesis within the sentence.

(2) Page 12, lines 422-423, could you consider providing a reference for the statement: "Trading networks involving this town may explain the presence of North African migrants so far north"?

Please note that these are only minor suggestions and there is no need for sending the manuscript for reviews again. Thank you!

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

In my opinion the authors have done a great job in responding to my comments, both in their careful rebuttal and in the high quality edits they have made to the manuscript. In particular I believe the edits to the figures supporting the relationship between dialectal areas and population structure add to its clarity and message. My concerns have been largely addressed and I believe the revised manuscript will be of great interest to readers of Nature Communications.

I have only one minor comment, which I leave to the authors and editors decision:
I think that the EEMS plot provided in the rebuttal to support the role of rivers in influencing population structure in France might be worth including in the supplementary materials for completeness. However I understand that fully discussing this analysis may be outside the scope of the manuscript.

Authors – The EEMS figure has been added to the Supplementary Figures and Figure S15. Two sentences have been added to the main text to discuss the results and a small section has been added on the Methods.
It reads now:

"Overlaps between regions of low migration and rivers also appear, although not systematically, in the effective migration surfaces retrieved with the EEMS[32] (Fig. S15). Specifically, we noted a significant decrease in migration across the Loire River (dark blue) and areas of low effective migration coincide with relatively large rivers such as Aulne (separating "Cornouaille" and "Léon") and Laïta (between "Cornouaille" and "Bretagne-Centre"), primarily in the westernmost part of Brittany. On the other hand, while the Vilaine River coincides with a region of low effective migration, its downstream part is in the centre of one cluster ("Guérande", close to the sea). However, it matches the North-West border of cluster "Nantes" where migration rates were inferred to be lower than expected. Recurring…"

Methods:

"Effective migration surface (EEMS) analysis

We estimated effective migration surfaces using the software EEMS[32]. The matrix of average pairwise genetic dissimilarities was generated for 82,362 SNPs (after prunning: --indep-pairwise 50 5 0.2) and 1414 individuals using the bed2diffs software included in the EEMS package. Samples were assigned to the nearest of 300 demes. We run ten independent MCMC chains, each with a random seed, for 10,000,000 iterations, including 9,900,000 burn-in iterations, thinning every 200 iterations. From the chain with the highest final log-likelihood, and we started a second round of ten EEMS chains using this chain as a starting point for 1,000,000 additional sampling iterations, thinning every 9,999 iterations. Log- posterior trace of ten replicate MCMC chains from the last round shows mixing and convergence of the independent EEMS runs. Plots were generated in R statistical software using the *rEEMSplots* package. "

I highly recommend accepting this article for publication.

Signed:
Ross Patrick Byrne

Reviewer #2 (Remarks to the Author):

This version is an improvement of an already excellent manuscript. The revised version reflects a thorough consideration of my previous feedback. I appreciate the attention given to my concerns and the comprehensive response provided by the authors.

I am happy with the level of detail of the analyses included in this second version of the manuscript. I believe the manuscript is now ready for publication in its current form and I have no further concerns to note.

Below, I highlight one potential typo and make one small suggestion. I would appreciate it if the authors could double-check these two potential issues, but I see no need to send the manuscript for another round of review as those two comments are only minor *suggestions*.

(1) Potential Typo? Page 11, lines 408-410 – While I couldn't pinpoint the exact issue, it seems there might be a missing word or parenthesis within the sentence.

Authors – Typo corrected .

(2) Page 12, lines 422-423, could you consider providing a reference for the statement: "Trading networks involving this town may explain the presence of North African migrants so far north"?

Authors – We aknowledge the reviewer for this remark. Indeed, our statement is very speculative. However, a recent study by Antonio et al 2024 eLife found a high proportion of genetic ancestry outliers and suggest that long range mobility was common in historical Europe. That said, we re-phrased to take into account these recent findinds. It reads as: “The exact reason for the presence of migrants of North African ancestry so far north is difficult to pinpoint. Nevertheless, it agrees with the recently reported high levels of mobility in historical Europe (51).”