This research titled "**Development and Characterization of SSR Markers in *Phoebe zhennan***"shows how important it is to produce SSR markers in order to protect and use Phoebe zhennan, an endangered tree species that is highly valued for its wood and ornamental value in China. The researchers were able to accurately measure the genetic diversity of 24 groups by finding more than 794,000 SSR sites and choosing 20 highly polymorphic markers. The results showed that there was a lot of variance among populations. For example, groups like SN and those in Sichuan and Hunan had high and stable genetic diversity, whereas populations like MT had reduced diversity. Frequent gene flow between locations also showed that populations were strongly connected.

The fact that P. zhennan has different genes from related species like P. bournei and P. sheareri is important since it shows that they evolved separately. These SSR markers are a great start for future species identification, germplasm conservation, and focused breeding initiatives. They are also a useful genomic resource because there isn't a whole reference genome yet.

Review on Abstract

The abstract makes a useful contribution to the conservation genetics of Phoebe zhennan by showing how SSR markers may be used to measure genetic variation in different populations. It does a good job of showing how important the species is to the environment and the economy, and it stresses how useful new molecular technologies are, especially when there isn't a reference genome. Quantitative indices like PIC, Shannon's index, and Nei's H are well explained, showing how strong and genetically diverse markers are, especially in some regional populations.

The abstract, on the other hand, has a lot of info in it and could need better organization and conciseness. Some important problems are inconsistent and the lack of a clear study goal, which makes it harder for the first reader to understand. Also, while it's easy to see why the focus is on Guizhou, it's not clear how the results can be used for the species as a whole. Adding a clear research goal, a clearer methodology, and a focus on the broader conservation implications would make the abstract clearer and more powerful resource when there isn't a complete reference genome.

Review on Introduction.

The introduction does a good job of showing that Phoebe zhennan is a tree species that is important for both the environment and the economy in China. The authors make a strong case for genetic conservation research by pointing out that the species is endangered (IUCN Red List, 1984), grows slowly, and faces ecological problems like overharvesting and habitat degradation. Adding the fact that frugivorous birds spread seeds gives the story more ecological depth. However, the ecological framing may be made stronger by adding references to demographic studies or studies on dispersal ecology to back up why this species is more vulnerable or genetically at risk. Adding the most recent conservation goals for endemic Lauraceae in China would make the story even better.

The authors say that the main research gap is the lack of genetic resources for P. zhennan. This is a statement that fits in well with the needs of forest tree genomics as a whole. It makes sense to move from EST-SSR methods to genomic SSRs, and the mention of previous works (Xiaodong et al., 2016; Qun et al., 2023) demonstrates that the author is well aware of how the field has changed.

Still, the reasons for transitioning from transcriptome-derived SSRs to genomic SSRs need to be made clearer, especially when it comes to their wider genome coverage and lack of bias compared to expressed regions. This would help make it clear why the current study is a methodological advance instead of just another step in the same direction.

The references are good and mostly up to current, however the synthesis could be better. Some of the changes are a little sudden, such when the text goes from talking about the benefits of SSR to taxonomy difficulties between species (lines 69–71) without a good link. Explicitly relating taxonomic uncertainties to the requirement for strong molecular markers could help bring things together better.

Also, sentence constructs like "Several molecular markers have proven to be more effective than other markers: DNA molecular markers..." are clumsy and would be clearer and flow better if they were grammatically fixed.

The last sentence ("Therefore, the development of molecular markers...") provides an excellent description of why the study was done. The beginning would be a lot stronger, though, if it clearly stated one or two research questions or hypotheses. This would let the reader know what to expect and make sure that the goals, techniques, and results are all in line with each other.

“markers:DNA", irregular spacing, and a stray quotation mark in "minor"differences" make the work look less professional. Acronyms like GD (genetic diversity) should be defined earlier and kept the same way. More focus on phrase economy and parallel construction would make the text easier to read.

The introduction sets up the ideas well with literature that is clear and relevant to the time. To make it more effective, the study might use more narrative coherence, better grammar, and clearer explanations of its unique contributions to molecular genetics and conservation biology. **These changes would make the introduction a good base for publication.**

Review on Material

- Getting 174 people from 24 populations in five provinces gives a lot of coverage and statistical power for studying genetic diversity and structure. This is a good thing about the study.

- But the language would be better if it explained clearly why the sample sizes were chosen for each group and why most of the locations were in Guizhou. Was this because it was easy to get to, it was important for the environment, or it was representative?

- Adding samples from other Phoebe species (P. bournei and P. chekiangensis) and Lindera megaphylla makes the taxonomic framework stronger. But it's not clear if these were only used to compare phylogenetic trees or if they helped make sure that the primers were specific during SSR validation.

The way the samples were collected and stored (using liquid nitrogen and ultra-low temperatures) is in line with best methods for keeping DNA intact. This makes people more sure about molecular work that comes after this.
There are a few small problems with the way the text is laid out and the way it looks. For example, "regionalChinese" should be written as "regional Chinese."
- The sign "-80#" looks like a formatting mistake; it probably means "–80 °C."
- Giving clear coordinates or elevation ranges for chosen locations would make the results more reproducible and give more ecological depth, especially for future landscape genomics studies.

This part sets up genetic analysis very well by showing that the samples were chosen carefully and that they cover a wide range of topics. Adding reasons for the sampling density, explicit goals for comparing species, and fixing small errors would make it clearer and more scientifically useful

Review on Method.

- It would be great to mention what the average yield or concentration range of DNA is for all the samples.
- For clarification, change the formatting problem "-80#" to "–80 °C."
- While the sequencing platform and service provider are mentioned, adding short data (such the total number of reads generated or the average coverage per sample) could make things clearer.
- It's great that technical replicates are used to keep genotyping errors to less than 1%, but it might be improved by looking at things like the type of replicate (biological vs. technical) and the downstream concordance rates.
- It's clear what the parameter limits for SSR identification are (for example, "≥6 dinucleotides"), but it would be more scientific if they explained why these cutoffs were chosen.
- Make sure to say whether SSR loci were filtered for repeat motif types (such not including AT-rich loci) or distribution bias before choosing primers.
- The symbols "¿L" and "p" are formatting errors and need to be changed to "µL" and the usual way to write volume.
- The criteria for choosing primers (such Tm and avoiding secondary structures) are good, however they may be made clearer by putting them in a table.
- Make it clear how missing data or unclear genotyping were handled (for example, by using 0/1 scoring or null alleles).
- Say whether the significance levels (for example, for Gst or AMOVA results) were changed for multiple comparisons or were based on permutation testing.
- Adding a flowchart or diagram of the whole analysis pipeline would make things clearer, especially for people who aren't used to SSR operations.

Your Methods section is well-written and pays good attention to DNA quality control, SSR library building, and data analysis. But it would be more rigorous if it fixed language and formatting mistakes, gave more reasons for sampling and parameter choices, and made clear how outgroups and replicates were treated.

- Clarity of Methodological Workflow: The order of the investigation makes sense, moving from scoring markers to using tools to look at variety and make guesses about population structure. But to make it easier to read, the description should use clearer transitions between the parts of the analysis. For instance, putting diversity measurements, cluster analysis, and population structure tools into their own paragraphs might make the flow better.
- Reproducibility and Precision: The "0 or 1" scoring method is discussed, but the rules for giving these ratings need to be made clearer (for example, should they be based on the presence or absence of peaks or the intensity of fluorescence?). This detail is very important for being able to repeat the experiment.
- It would be good to indicate what model and parameters were used in the STRUCTURE

analysis, such as the number of runs, the length of the burn-in period, and the number of MCMC iterations.
- For AMOVA, say how many permutations were utilized to check for significance.
- Software Citation Consistency: Even though the names and versions of the software are given, it's best to mention the original articles or developers for each tool so that people may find them again in the future. For instance, utilize Peakall & Smouse for GenAlEx and Excoffier et al. for Arlequin.
- Suggestions for formatting:
- Think about using italics for software names or formatting that depends on the journal style.
- Use the same versioning notation for all versions (for example, "v3.25" vs. "3.5.2.2"—either all with "v" prefix or none).
- Integration of visual aids: It's good that the iTOL online portal was brought up, but it would be better if you briefly said what kind of visual output you got (for example, a dendrogram with bootstrap values?).

Review on Result

Section 3.1: Characterization of SSR Loci

The paragraph does a good job of giving important descriptive statistics about the distribution of SSR. The content is hard to read, though. It would be easier logically arranging statistics (for example, by motif length or repetition type) and breaking out lengthy phrases would make it easier to understand.

- It is evident that dinucleotide SSRs are the most common type, but you might want to talk about whether this pattern fits with the genomic features of P. zhennan or with what has been found in other species that are similar.
- The phrase "the number of SSR loci decreased gradually with a rise in repeat numbers" (lines 174–175) gives some information, but it would be better to back up the claim with data (for example, a percentage drop across categories).
- Consistency and Terminology: Make sure that the terms you use are always the same. For example, only use "mononucleotide SSRs" when it's necessary to avoid repeating yourself.
- Be careful about the space around percentages. For example, "84.73 %" should be "84.73%" to meet standard style.

- Use of Figures and Tables: The text talks about Table 2 and Figures 1A and 1B, but it would be beneficial to have a short comment explaining what each figure or table shows. For example, saying that Figure 1A is "a graphical breakdown of SSR motif types" and Figure 1B is "distribution of repeat counts" could help the reader understand what they are looking at.
- If you can, it would be great to include the total SSR density (the number of SSRs per kb) to give the abundance some context.

- In a scientific context: You could want to briefly talk about what the A/T dominance in mononucleotide repeats or the AG/CT abundance in dinucleotide SSRs means for biology or function.

Part 3.2: Looking at the SSR Markers' Polymorphism
- Planning the experiment and evaluating the markers:
- The selection of multi-phase markers (gel electrophoresis followed by capillary screening) is well-organized. But the definitions of "amplification efficiency" and "specificity" are still a little unclear. Adding quantitative thresholds or scoring details would make things more reproducible and clear.
- Think about making it clear whether the 108 primer pairs were made from scratch using the SSR-enriched library or used from other research.
- Presenting and Interpreting Data:
- The PIC range (0.777–0.903) and mean (0.859) convincingly back up the allegation of high polymorphism. However, a short comparison with similar studies in Phoebe or Lauraceae species would help make the direct conclusion that these SSRs are good for GD investigations.
- The small difference between the observed heterozygosity (Ho = 0.705) and the anticipated heterozygosity (He = 0.487) needs more explanation. Does this mean that there is selection, a heterozygote advantage, or a population structure?
- Technical Edits: There seems to be a sentence cut off on line 200: "assessing the GD of (Table 3)." Please finish or change for clarity.
- Use the same way to write the gene diversity index (H) across all findings. For example, always use "Nei's gene diversity (H)."

Section 3.3: Genetic Diversity (GD) Analysis of the Guizhou Population—Statistical Strength and Comparison Analysis of Genetic Diversity (GD) in the Guizhou Population: Statistical Robustness and Comparisons
- The patterns of GD within the region are shown in detail. But you would want to say whether the differences in Na, Ho, He, etc. were statistically especially when comparing SN to MT or other provinces.
- The conclusion concerning inbreeding (Ho < He) makes sense, however it would be good to have F_IS values to back it up with numbers.
- Cross-Regional Comparisons: The comparison with BJ, GL, ZG, and TN populations offers geographic breadth, but it doesn't give enough information about how the samples were balanced and how big they were in each region, which could affect the GD estimates.
- Adding other Phoebe species gives us a really interesting view of how different species interact with each other. Still, it would make the case stronger to make it clear if these were taken from areas with similar ecosystems or from protected areas.
- Interpretive Insight: - The SN population has a greater GD, however it would be beneficial to briefly think about biological or human-made causes that might be making this richness (for example, habitat fragmentation, protection Suggestions for the following:

- Visual Enhancement: Use summary visual aids like a heatmap of GD indices across populations or scatter plots of Ho vs. He to quickly show patterns of inbreeding or GD hotspots.
- Consistency in language: Make sure that all statistical terminology and abbreviations (such Ho, He, H, and I) are formatted the same way and that each one is defined the first time it is used so that readers from different fields may understand them.

Comments from the reviewer on sections 3.4 to 3.6: Genetic Differentiation, Structure, and Clustering

- Clear and Defined:
- The explanations of Fis, Fit, and Fst are useful, but the reasoning on line 234 doesn't seem to make sense: inbreeding usually makes more homozygotes, not "pure heterozygotes." This has to be changed to be more accurate in terms of biology.
- To prevent repeating ideas, such defining Fis and Fit, they might be condensed.
- Data Interpretation: The Fst values that were reported (mean = 0.175) show that there is moderate genetic differentiation. Think about saying clearly whether this amount is statistically significant, and then compare it to what you would expect from similar studies in Phoebe or other Lauraceae species.
- The estimates of gene flow ($Nm > 1$) are clear, but more has to be said about how they relate to population structure and isolation by distance.
- AMOVA Results: - The AMOVA output is clearly shown. It is important to protect local genetic diversity because 90.57% of variation happens within populations.
- Saying how many permutations were used in AMOVA (like 10,000) would make it easier to repeat the process.

3.5 Analysis of Population Structure—STRUCTURE Analysis and Q Value Thresholding:
- The STRUCTURE results are easy to read, with $Q \geq 0.6$ as a grouping rule. However, mentioning admixture or different K-values (the ΔK technique) would make it more likely that $K = 3$ is the right choice.
- Think about making it clear how people with $Q < 0.6$ were treated. Were they put in a new group, mixed in with others, or left out?
- Genetic Pool Interpretation: - The fact that P. sheareri, P. chekiangensis, and P. bournei are grouped together in one subgroup shows that they have diverged from one other. This section would be better if it included a short discussion of possible hybridization or shared evolutionary roots.
- Comment on the visualization:
- The overview of Fig. 2 and Fig. 3 is helpful but not very detailed. Think about adding some interpretive detail, such what does the overlap of red, green, and blue groupings in space mean for ecology or evolution?

3.6 Cluster and PCA Analyses - UPGMA Clustering: - The results of the clustering are mostly correct. It's interesting that the categorization went from 3 groups to 5. You might want to mention the linkage threshold used in clustering or bootstrap values if you have them.
- The connection between P. bournei, P. sheareri, and P. chekiangensis is strong and could mean that they are closely related in terms of their evolutionary history, but we should be careful without more sequence data.
- - Discussion on Outlier and Species Divergence: It's vital to note that the Yanhe County sample may be a different species. This point has to be made more clear, maybe as a way to lead into a recommendation for taxonomy or conservation.
- PCA Meaning: It's good that the PCA results mostly match up with the cluster analysis. Make sure to say how much of the variance PC1 and PC2 explain (for example, % of the total variance), and talk about what might be causing the cross-pattern among P. zhennan individuals.

- It's crucial to note that the Yanhe County sample may be a different species. This point needs greater attention.

Review on Discussion

- Clarity and logical consistency:
- Line 331: The phrase "Ho value of 0.701 was greater than the expected He value of 0.737" seems to contradict itself. Please verify again, as Ho < He in this statement. The values or the way they are understood should be fixed.
- Line 295–296: It's possible that the phrase "these two P. zhennan species" is a mistake. Maybe it was meant to say P. zhennan and P. bournei?
- Interpretative Depth: - It is said that SSR motif types can be used to make evolutionary inferences (lines 314–315), however these assertions should be taken with a grain of salt unless they are backed up by functional genomic evidence.
- It could be questionable to say that mononucleotide and dinucleotide repetitions show a "high evolutionary level" (line 321). It might be safer to say that this is a theorized correlation rather than a conclusion.
- Structure and Readability: Some parts (like lines 331–357) have extensive chains of values with very little story. You could want to put them all together in a table or heatmap and then use text to point out the most important points.
- The phrase "Genetic diversity is an important index..." (line 322) should be better used earlier to explain why it is vital to look at GD in P. zhennan conservation.

- Reducing Redundancy: - There is some overlap between the interpretation of earlier results (e.g., population-level GD) and the discussion here. To make the text flow better, cut down on the repetitious listing of population GD numbers and put more emphasis on ecological or historical causes.

- Policy and Practical Implications: - The debate talks about in situ and ex situ conservation, but it might go into more detail on specific conservation recommendations, like: - Whether certain populations should be prioritized as seed sources
- The possibility of making genetic corridors to stop fragmentation - Working with national or regional conservation policy frameworks

Minor Suggestions: - Make sure that the formatting of taxonomic information is the same (always italicize species names).
- Make sure you use "GD index (H)" and "He" and "Ho" consistently and correctly. Making the difference clear can help readers who aren't as knowledgeable with population genetics not get confused.
- When talking about significant findings, directly name the tools used (such MISA or Arlequin) to make the methods more clear.This area has a lot of analytical and conservation value overall. Improving the story's logic and making the ecological interpretation more in-depth will make the manuscript's contribution to molecular conservation studies of subtropical tree species - Scope and Depth: The conclusions might be clearer about how the results affect breeding techniques, conservation priorities, or future efforts to clarify taxonomic issues.
- Think about ways to make the study more interesting, including whether this is the first time that SSR characterisation has been done on a broad scale in P. zhennan using this method, or how it is different from other studies (for example, those that used EST-derived markers or SNP-based studies).

- Consistency with Results: The claim that "genetic differentiation being primarily intra-population" (line 435) might be a little misleading because AMOVA indicated that most of the variance was within populations, but the genetic differentiation (Fst) still showed substantial divergence. Rewording could stop people from misunderstanding.
- Terminological Precision: "Effectively distinguished" (line 438) is a little imprecise. Does this mean STRUCTURE analysis, clustering, or allele specificity? Please tell us what approach or proof you employed to back up this assertion of interspecific differentiation.
- Clarity and Flow: To make it easier to read, think about breaking up large words into shorter, more explicit ideas. This is especially important when going from GD patterns to interpopulation interactions and interspecific resolution.

This work produced SSR molecular markers using a genomic library of P. zhennan that has a lot of SSRs in it. Using 20 highly polymorphic SSR loci, we found that all of the Guizhou populations had a lot of genetic diversity. The SN population had the most diversity, while the MT population had the least. Most of the genetic variance was discovered within populations, although there was also some mild genetic difference between groups.
Cluster analysis showed that the populations from Zigong (ZG), Tongnan (TN), and Tongzi (TZ) are very closely related genetically. The polymorphic SSR markers also successfully separated P.

zhennan from closely related species including P. bournei, P. sheareri, and P. chekiangensis. This shows how useful they are for genetic studies within and between species.
These results give genetic monitoring, breeding programs, and conservation planning in the Phoebe genus useful molecular resources.

- Scope and Depth: The conclusions might be clearer about how the results affect breeding techniques, conservation priorities, or future efforts to clarify taxonomic issues.
- Think about ways to make the study more interesting, including whether this is the first time that SSR characterisation has been done on a broad scale in P. zhennan using this method, or how it is different from other studies (for example, those that used EST-derived markers or SNP-based studies).

- Consistency with Results: The claim that "genetic differentiation being primarily intra-population" (line 435) might be a little misleading because AMOVA indicated that most of the variance was within populations, but the genetic differentiation (Fst) still showed substantial divergence. Rewording could stop people from misunderstanding.
- Terminological Precision: "Effectively distinguished" (line 438) is a little imprecise. Does this mean STRUCTURE analysis, clustering, or allele specificity? Please tell us what approach or proof you employed to back up this assertion of interspecific differentiation.
- Clarity and Flow: To make it easier to read, think about breaking up large words into shorter, more explicit ideas. This is especially important when going from GD patterns to interpopulation interactions and interspecific resolution.
Suggested Change (for clarity and polish):

Suggested Change
This work produced SSR molecular markers using a genomic library of P. zhennan that has a lot of SSRs in it. Using 20 highly polymorphic SSR loci, we found that all of the Guizhou populations had a lot of genetic diversity. The SN population had the most diversity, while the MT population had the least. Most of the genetic variance was discovered within populations, although there was also some mild genetic difference between groups.
Cluster analysis showed that the populations from Zigong (ZG), Tongnan (TN), and Tongzi (TZ) are very closely related genetically. The polymorphic SSR markers also successfully separated P. zhennan from closely related species including P. bournei, P. sheareri, and P. chekiangensis. This shows how useful they are for genetic studies within and between species.
These results give genetic monitoring, breeding programs, and conservation planning in the Phoebe genus useful molecular resources.

Review on References.

- Consistent formatting: Use the same style for all authors (for example, put initials after last names and spaces after commas).
- For example, "Isa0 Rita MR" (line 444) probably has an OCR error in "Isa0."
- "Hanbo Y, Wenna A..." (line 471) or "Christian S." (line 453): pick one style.
- Use italics for journal names and sentence case for article titles, unless the style of the journal is different.
- Ensure all DOIs are correctly formatted and prefixed with https://doi.org/ for hyperlinking.
- Look for typos and duplicates:
- You might want to check to see if line 279 mistakenly says "P. bournei, P. chekiangensis, and P. sheareri" over and over again.
- Some articles, including Changming et al. 2015 and Manosh Kumar et al. 2020, say "0-0" for page numbers. These probably need to be fixed or replaced with placeholders.
- Sources that haven't been published:
- Clearly mark theses and dissertations as such and check their DOI or repository URL, like Jian (2016) did.
- Make sure that conference abstracts or data-only records (if they exist) are labeled correctly.
- Citation Matching and Alphabetical Sorting: - If this reference list is meant to be an alphabetical bibliography, please check the sorting.
- Make sure that each citation in the manuscript is correct (and the other way around).
- Checks for Consistency:
- Make sure that the citation style is the same throughout, especially for initials, italics for journal names, and sentence casing for article titles.

- Make sure that page ranges are the same (for example, don't use "0-0") and that missing page numbers or volumes are complete where possible.
- Make sure that DOIs are formatted correctly with the right URLs (i.e., all start with https://doi.org/).
- Alphabetical Ordering: If your target journal (like PeerJ) requires it, make sure that all entries are in the right order based on the last name of the first author.
- Cross-Referencing: - Make sure that all of the in-text citations match the entries here and that there are no references that are cited but not included or the other way around.
Read the authors' advice for how to properly prepare the references if you want to send this to PeerJ.