A Manuscript Microscope 'Sentence Audit' of:

MultiGWAS: An integrative tool for Genome Wide Association Studies (GWAS) in tetraploid organisms

What is the Manuscript Microscope Sentence Audit?

The Manuscript Microscope Sentence Audit is a research paper introspection system that parses the text of your manuscript into minimal sentence components for faster, more accurate, enhanced proofreading.

Why use a Sentence Audit to proofread your manuscript?

- Accelerated Proofreading: Examine long technical texts in a fraction of the usual time.
- Superior Proofreading: Detect subtle errors that are invisible to traditional methods.
- Focused Proofreading: Inspect each individual sentence component in isolation.
- Reliable Proofreading: Ensure every single word of your manuscript is correct.
- Easier Proofreading: Take the hardship out of crafting academic papers.

Bonus 1: Improved Productivity: Rapidly refine rough drafts to polished papers.
Bonus 2: Improved Authorship: Cultivate a clear, concise, consistent, writing style.
Bonus 3: Improved Reputation: Become known for rigorously precise publications.

Features of the Sentence Audit:

The Sentence Audit combines two complementary proofreading approaches:

1. Each sentence of your text is parsed and displayed in isolation for focused inspection.
2. Each individual sentence is further parsed into Minimal Sentence Components for a deeper review of the clarity, composition and consistency of the language you used.

The Minimal Sentence Components shown are the smallest coherent elements of each sentence of your text as derived from it's conjunctions, prepositions and selected punctuation symbols (i.e. commas, semicolons, round and square brackets).

The combined approaches ensure easier, faster, more effective proofreading.

Comments and Caveats:

• The sentence parsing is achieved using a prototype natural language processing pipeline written in Python and may include occasional errors in sentence segmentation.

• Depending on the source of the input text, the Sentence Audit may contain occasional html artefacts that are parsed as sentences (E.g. "Download figure. Open in new tab").

• Always consult the original research paper as the true reference source for the text.

Contact Information:

To get a Manuscript Microscope Sentence Audit of any other research paper, simply forward any copy of the text to John.James@OxfordResearchServices.com.

All queries, feedback or suggestions are also very welcome.

Research Paper Sections:

The sections of the research paper input text parsed in this audit.

Section No.	Headings	Sentences
Section: 1	Abstract	15
Section: 2	1 Introduction	13
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0

Title	**MultiGWAS: An integrative tool for Genome Wide Association Studies (GWAS) in tetraploid organisms**

S1 [001]   Abstract

S1 [002]   Summary

Summary

S1 [003]   The Genome-Wide Association Studies (GWAS) are essential to determine the genetic bases of either ecological or economic phenotypic variation across individuals within populations of the model and non-model organisms.

The Genome-Wide Association Studies ...

... (GWAS) ...

... are essential ...

... to determine the genetic bases ...

... of either ecological ...

... or economic phenotypic variation ...

... across individuals ...

... within populations ...

... of the model ...

... and non-model organisms.

S1 [004]   For this research question, the GWAS replication testing different parameters and models to validate the results' reproducibility is common.

For this research question, ...

... the GWAS replication testing different parameters ...

... and models ...

... to validate the results' reproducibility is common.

S1 [005]   However, straightforward methodologies that manage both replication and tetraploid data are still missing.

However, ...

... straightforward methodologies ...

... that manage both replication ...

... and tetraploid data are still missing.

S1 [006]   To solve this problem, we designed the MultiGWAS, a tool that does GWAS for diploid and tetraploid organisms by executing in parallel four software, two designed for polyploid data (GWASpoly and SHEsis) and two for diploids data (GAPIT and TASSEL).

To solve this problem, ...

... we designed the MultiGWAS, ...

... a tool ...

... that does GWAS ...

... for diploid ...

... and tetraploid organisms ...

... by executing ...

... in parallel four software, ...

... two designed ...

... for polyploid data ...

... (GWASpoly ...

... and SHEsis) ...

... and two ...

... for diploids data ...

... (GAPIT ...

... and TASSEL).

S1 [007] MultiGWAS has several advantages.

MultiGWAS has several advantages.

S1 [008] It runs either in the command line or in a graphical interface; it manages different genotype formats, including VCF.

It runs either ...

... in the command line ...

... or in a graphical interface; ...

... it manages different genotype formats, ...

... including VCF.

S1 [009] Moreover, it allows control for population structure, relatedness, and several quality control checks on genotype data.

Moreover, ...

... it allows control ...

... for population structure, ...

... relatedness, ...

... and several quality control checks ...

... on genotype data.

S1 [010] Besides, MultiGWAS can test for additive and dominant gene action models, and through a proprietary scoring function, select the best model to report its associations.

Besides, ...

... MultiGWAS can test ...

... for additive ...

... and dominant gene action models, ...

... and ...

... through a proprietary scoring function, ...

... select the best model ...

... to report its associations.

S1 [011] Finally, it generates several reports that facilitate identifying false associations from both the significant and the best-ranked association SNP among the four software.

Finally, ...

... it generates several reports ...

... that facilitate identifying false associations ...

... from both the significant ...

... and the best-ranked association SNP ...

... among the four software.

S1 [012] We tested MultiGWAS with public tetraploid potato data for tuber shape and several simulated data under both additive and dominant models.

> We tested MultiGWAS ...
>
> ... with public tetraploid potato data ...
>
> ... for tuber shape ...
>
> ... and several simulated data ...
>
> ... under both additive ...
>
> ... and dominant models.

S1 [013] These tests demonstrated that MultiGWAS is better at detecting reliable associations than using each of the four software individually.

> These tests demonstrated ...
>
> ... that MultiGWAS is better ...
>
> ... at detecting reliable associations ...
>
> ... than ...
>
> ... using each ...
>
> ... of the four software individually.

S1 [014] Moreover, the parallel analysis of polyploid and diploid software that only offers Multi-GWAS demonstrates its utility in understanding the best genetic model behind the SNP association in tetraploid organisms.

> Moreover, ...
>
> ... the parallel analysis ...
>
> ... of polyploid ...
>
> ... and diploid software ...
>
> ... that ...
>
> ... only offers Multi-GWAS demonstrates its utility ...
>
> ... in understanding the best genetic model behind the SNP association ...
>
> ... in tetraploid organisms.

S1 [015] Therefore, MultiG-WAS probed to be an excellent alternative for wrapping GWAS replication in diploid and tetraploid organisms in a single analysis environment.

> Therefore, ...
>
> ... MultiG-WAS probed ...
>
> ... to be an excellent alternative ...
>
> ... for wrapping GWAS replication ...
>
> ... in diploid ...
>
> ... and tetraploid organisms ...
>
> ... in a single analysis environment.

S2 [016] **1 Introduction**

S2 [017] The Genome-Wide Association Studies (GWAS) comprise statistical tests that identify which variants through the whole genome of a large number of individuals are associated with a specific trait [Cantor et al., 2010, Begum et al., 2012].

> The Genome-Wide Association Studies ...
>
> ... (GWAS) ...
>
> ... comprise statistical tests ...
>
> ... that identify ...
>
> ... which variants ...
>
> ... through the whole genome ...
>
> ... of a large number ...
>
> ... of individuals are associated ...
>
> ... with a specific trait ...
>
> ... [Cantor et al., 2010, ...
>
> ... Begum et al., 2012].

S2 [018] This methodology started with humans and several model plants, such as rice, maize, and Arabidopsis [Lauc et al., 2010, Tian et al., 2011, Cao et al., 2011, Korte and Farlow, 2013, Han and Huang, 2013].

> This methodology started ...
>
> ... with humans ...
>
> ... and several model plants, ...
>
> ... such as rice, ...
>
> ... maize, ...
>
> ... and Arabidopsis ...
>
> ... [Lauc et al., 2010, ...
>
> ... Tian et al., 2011, ...
>
> ... Cao et al., 2011, ...
>
> ... Korte ...
>
> ... and Farlow, 2013, ...
>
> ... Han ...
>
> ... and Huang, 2013].

S2 [019] Because of the advances in the high-throughput sequencing technology and the decline of the sequencing cost in recent years, there is an increase in the availability of genome sequences of different organisms at a faster rate [Ekblom and Galindo, 2011, Ellegren, 2014].

> Because ...
>
> ... of the advances ...
>
> ... in the high-throughput sequencing technology ...
>
> ... and the decline ...
>
> ... of the sequencing cost ...
>
> ... in recent years, ...
>
> ... there is an increase ...
>
> ... in the availability ...
>
> ... of genome sequences ...
>
> ... of different organisms ...
>
> ... at a faster rate ...
>
> ... [Ekblom ...
>
> ... and Galindo, 2011, ...
>
> ... Ellegren, 2014].

End of Sample Audit