Predicting functional consequences of mutations using molecular interaction network features

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.

Manuscript Source: https://www.biorxiv.org/content/10.1101/2021.03.05.433991v1
Manuscript Authors: Kivilcim Ozturk & Hannah Carter

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	10
Section: 2	INTRODUCTION	17
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	**Predicting functional consequences of mutations using molecular interaction network features**

S1 [001] Abstract

S1 [002] Variant interpretation remains a central challenge for precision medicine.

> Variant interpretation remains a central challenge ...
>
> ... for precision medicine.

S1 [003] Missense variants are particularly difficult to understand as they change only a single amino acid in protein sequence yet can have large and varied effects on protein activity.

> Missense variants are particularly difficult ...
>
> ... to understand ...
>
> ... as they change ...
>
> ... only a single amino acid ...
>
> ... in protein sequence ...
>
> ... yet can have large ...
>
> ... and varied effects ...
>
> ... on protein activity.

S1 [004] Numerous tools have been developed to identify missense variants with putative disease consequences from protein sequence and structure.

> Numerous tools have been developed ...
>
> ... to identify missense variants ...
>
> ... with putative disease consequences ...
>
> ... from protein sequence ...
>
> ... and structure.

S1 [005] However, biological function arises through higher order interactions among proteins and molecules within cells.

> However, ...
>
> ... biological function arises ...
>
> ... through higher order interactions ...
>
> ... among proteins ...
>
> ... and molecules ...
>
> ... within cells.

S1 [006] We therefore sought to capture information about the potential of missense mutations to perturb protein interaction networks by integrating protein structure and interaction data.

> We therefore sought ...
>
> ... to capture information ...
>
> ... about the potential ...
>
> ... of missense mutations ...
>
> ... to perturb protein interaction networks ...

... by integrating protein structure ...

... and interaction data.

S1 [007]   We developed 16 network-based annotations for missense mutations that provide orthogonal information to features classically used to prioritize variants.

We developed 16 network-based annotations ...

... for missense mutations ...

... that provide orthogonal information ...

... to features classically used ...

... to prioritize variants.

S1 [008]   We then evaluated them in the context of a proven machine-learning framework for variant effect prediction across multiple benchmark datasets to demonstrate their potential to improve variant classification.

We then evaluated them ...

... in the context ...

... of a proven machine-learning framework ...

... for variant effect prediction ...

... across multiple benchmark datasets ...

... to demonstrate their potential ...

... to improve variant classification.

S1 [009]   Interestingly, network features resulted in larger performance gains for classifying somatic mutations than for germline variants, possibly due to different constraints on what mutations are tolerated at the cellular versus organismal level.

Interestingly, ...

... network features resulted ...

... in larger performance gains ...

... for classifying somatic mutations ...

... than ...

... for germline variants, ...

... possibly ...

... due to different constraints ...

... on what mutations are tolerated ...

... at the cellular versus organismal level.

S1 [010]   Our results suggest that modeling variant potential to perturb context-specific interactome networks is a fruitful strategy to advance in silico variant effect prediction.

Our results suggest ...

... that modeling variant potential ...

... to perturb context-specific interactome networks is a fruitful strategy ...

... to advance ...

... in silico variant effect prediction.

S2 [011]   **INTRODUCTION**

S2 [012] Advances in high throughput sequencing technologies have resulted in the rapid accumulation of genomic data and allowed profiling of patient genomes in clinical settings.

> Advances ...
>
> ... in high throughput sequencing technologies have resulted ...
>
> ... in the rapid accumulation ...
>
> ... of genomic data ...
>
> ... and allowed profiling ...
>
> ... of patient genomes ...
>
> ... in clinical settings.

S2 [013] Such studies frequently uncover previously unobserved and uncharacterized genetic variants of ambiguous relevance to health, making variant interpretation an important challenge in precision medicine [1].

> Such studies frequently uncover previously unobserved ...
>
> ... and uncharacterized genetic variants ...
>
> ... of ambiguous relevance ...
>
> ... to health, ...
>
> ... making variant interpretation an important challenge ...
>
> ... in precision medicine ...
>
> ... [1].

S2 [014] Missense mutations are particularly challenging as they only change a single amino acid in a protein sequence yet can have effects spanning no difference to complete loss of function.

> Missense mutations are particularly challenging ...
>
> ... as they ...
>
> ... only change a single amino acid ...
>
> ... in a protein sequence ...
>
> ... yet can have effects spanning no difference ...
>
> ... to complete loss ...
>
> ... of function.

S2 [015] Numerous methods have been developed to prioritize functional missense variants [2–10].

> Numerous methods have been developed ...
>
> ... to prioritize functional missense variants ...
>
> ... [2–10].

S2 [016] Typically, these tools rely on protein sequence/structure information to predict variant effects at the protein level, and the scores they provide tend to capture coarse grained estimates of impact (e.g damaging, benign, tolerated).

> Typically, ...
>
> ... these tools rely ...
>
> ... on protein sequence/structure information ...
>
> ... to predict variant effects ...
>
> ... at the protein level, ...
>
> ... and the scores they provide tend ...
>
> ... to capture coarse grained estimates ...
>
> ... of impact ...
>
> ... (e.g damaging, ...

... benign, ...

... tolerated).

S2 [017]   Biological functions and cellular behaviors arise from interactions among proteins and other molecules within cells, and biological systems evolve to be robust to random error [11].

Biological functions ...

... and cellular behaviors arise ...

... from interactions ...

... among proteins ...

... and other molecules ...

... within cells, ...

... and biological systems evolve ...

... to be robust ...

... to random error ...

... [11].

S2 [018]   Diseases are often associated with perturbations to protein interactions, different perturbations can result in different phenotypes [12], and the level of impact caused by mutations to the underlying molecular interaction network may determine the likelihood of generating a phenotype [13].

Diseases are often associated ...

... with perturbations ...

... to protein interactions, ...

... different perturbations can result ...

... in different phenotypes ...

... [12], ...

... and the level ...

... of impact caused ...

... by mutations ...

... to the underlying molecular interaction network may determine the likelihood ...

... of generating a phenotype ...

... [13].

S2 [019]   For example, loss of function mutations were more likely to be tolerated when they affected proteins at the periphery of the interactome [14].

For example, ...

... loss ...

... of function mutations were more likely ...

... to be tolerated ...

... when they affected proteins ...

... at the periphery ...

... of the interactome ...

... [14].

S2 [020]   Similarly, variants that otherwise were predicted to have little effect were more likely to be deleterious if they had a large number of interaction partners [15].

Similarly, ...

... variants ...

... that otherwise were predicted ...

... to have little effect were more likely ...

End of Sample Audit