

# **The Little Book of Smiley Plots**

**A collection of ancient DNA patterns and their causes**

The SPAAM Community

2025-01-05

# Table of contents

<b>Preface</b>	<b>3</b>
<b>Introduction</b>	<b>4</b>
What are deamination patterns? . . . . .	4
How to generate damage patterns . . . . .	4
Genomics . . . . .	4
Metagenomics . . . . .	4
 <b>I   Valid Smiley Plots</b>	 <b>5</b>
<b>1   Smiley Plot of Double Stranded DNA Libraries</b>	<b>6</b>
<b>2   Smiley Plot of Single Stranded DNA Libraries</b>	<b>7</b>
 <b>II   Problematic Smiley Plots</b>	 <b>8</b>
<b>3   Smiley Plot of Insufficient Reads</b>	<b>9</b>
<b>References</b>	<b>10</b>

# Preface

A key part of any ancient DNA project is to show that the DNA is exactly that - ancient. A key authentication method is to show the presence of elevated C to T deamination patterns (and the complementary G to A) at the end of DNA molecules, originally reported by (Briggs et al. 2007).

These patterns can be plotted in what have been colloquially known as 'Smiley Plots. However, there can be a wide range of smiley plots, some which show valid ancient DNA, and others that do not - either due to not actually having true ancient DNA but also from laboratory and/or bioinformatic artifacts.

This book aims

# Introduction

## What are deamination patterns?

## How to generate damage patterns

There is a range of software that can generate damage pattern plots from ancient DNA NGS libraries. The vast majority of tools require to be of sequencing reads aligned to a reference genome. Here we make suggestions of some tools that you can use to generate such plots.

### Genomics

These tools general take BAM files as input (i.e., after mapping of FASTQ files to a reference genome using a short-read aligner )

- mapDamage
  - Source: <https://github.com/ginolhac/mapDamage>
  - Documentation: <https://ginolhac.github.io/mapDamage>
  - Citation: (Jónsson et al. 2013)
- DamageProfiler
  - Source: <https://github.com/Integrative-Transcriptomics/DamageProfiler>
  - Documentation: <https://damageprofiler.readthedocs.io/en/latest/>
  - Citation: (Neukamm, Peltzer, and Nieselt 2021)

### Metagenomics

- MaltExtract
  - Source: <https://github.com/rhuebler/MaltExtract>
  - Documentation: <https://github.com/rhuebler/MaltExtract>
  - Citation: (Hübler et al. 2019)

**Part I**

**Valid Smiley Plots**

# 1 Smiley Plot of Double Stranded DNA Libraries

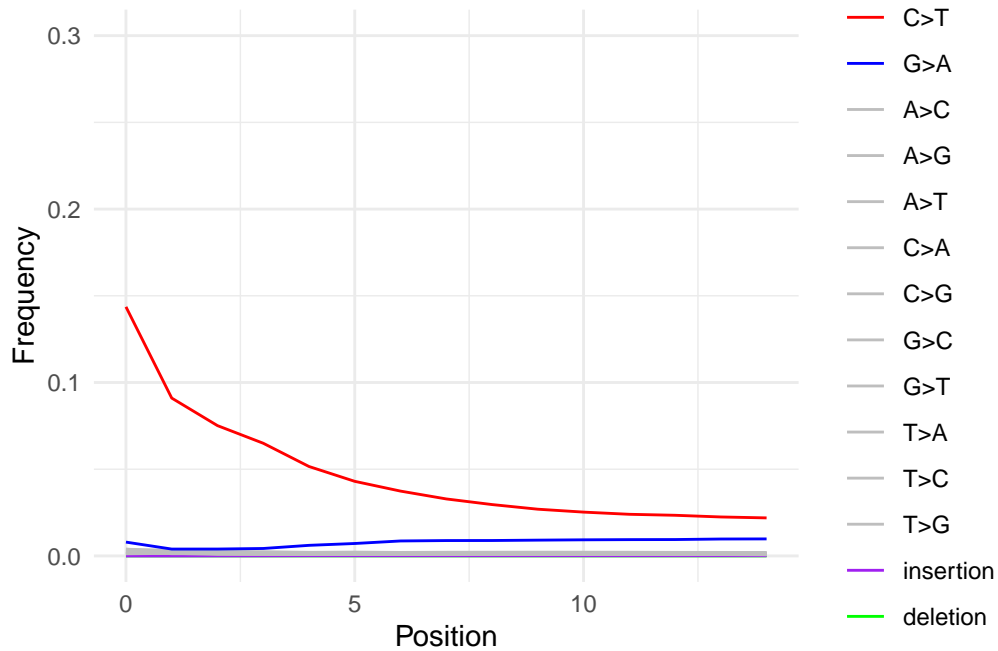


Figure 1.1: Example of a smiley plot of a double stranded DNA library. Data taken from (Star et al. 2017). Damage data calculated using [DamageProfiler](#) and plotted using R and ggplot2.

This is the ‘classical’ ancient DNA plot that you will see most often. You expect to see a smooth curve from the beginning of the read (position 1) to a flat line in the middle (e.g. positions 10-25 in mapDamage plots).

If you get such a plot in ancient DNA double-stranded libraries, this is a good indication you have authentic ancient DNA!

## **2 Smiley Plot of Single Stranded DNA Libraries**

**Part II**

**Problematic Smiley Plots**



### **3 Smiley Plot of Insufficient Reads**

# References

- Briggs, Adrian W, Udo Stenzel, Philip L F Johnson, Richard E Green, Janet Kelso, Kay Prüfer, Matthias Meyer, et al. 2007. “Patterns of Damage in Genomic DNA Sequences from a Neandertal.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (37): 14616–21. <https://doi.org/10.1073/pnas.0704665104>.
- Hübner, Ron, Felix M Key, Christina Warinner, Kirsten I Bos, Johannes Krause, and Alexander Herbig. 2019. “HOPS: Automated Detection and Authentication of Pathogen DNA in Archaeological Remains.” *Genome Biology* 20 (1): 280. <https://doi.org/10.1186/s13059-019-1903-0>.
- Jónsson, Hákon, Aurélien Ginolhac, Mikkel Schubert, Philip L F Johnson, and Ludovic Orlando. 2013. “mapDamage2.0: Fast Approximate Bayesian Estimates of Ancient DNA Damage Parameters.” *Bioinformatics* 29 (13): 1682–84. <https://doi.org/10.1093/bioinformatics/btt193>.
- Neukamm, Judith, Alexander Peltzer, and Kay Nieselt. 2021. “DamageProfiler: Fast Damage Pattern Calculation for Ancient DNA.” *Bioinformatics* 37 (20): 3652–53. <https://doi.org/10.1093/bioinformatics/btab190>.
- Star, Bastiaan, Sanne Boessenkool, Agata T Gondek, Elena A Nikulina, Anne Karin Hufthammer, Christophe Pampoulie, Halvor Knutsen, et al. 2017. “Ancient DNA Reveals the Arctic Origin of Viking Age Cod from Haithabu, Germany.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (34): 9152–57. <https://doi.org/10.1073/pnas.1710186114>.