



LATEX Workshop

Thijs Janzen



What is L^AT_EX ?

- TeX (from Greek τέχνη, ‘art’ or ‘craft’) language developed for consistent typesetting in 1978 by Daniel Knuth
 - Goals:
 - 1) to allow anybody to produce high-quality books with minimal effort
 - 2) provide a system that would give exactly the same results on all computers, at any point in time
- Leslie Lamport developed macros in the early 1980’s to make writing TeX easier: L^ATeX
- Currently, TeX is at version 3.141592653

What is L^AT_EX ?

- Widely used across academia, mainly in mathematics, computer science, economics, engineering and physics
- But also by book and journal publishers (Elsevier, Cambridge University Press, Oxford University Press etc)
- Separate presentation from content
- Use commands for makeup, instead of WYSIWYG
 - ‘Programming for Writing’
 - A bit similar in feel to R and RMarkdown
- Version control using git and GitHub
- Free!!

Examples

Haplotype block dynamics in hybrid populations

Thijs Janzen^{*,1}, Arne Nolte^{*} and Arne Traulsen^{*}

^{*}Max-Planck-Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306, Plön, Germany

ABSTRACT When two species hybridize, the genome of the hybrid offspring becomes a blend of the genome of either of the ancestral species. Over time, due to recombination, blocks appear in the genome of the hybrids that contain genomic material from either of the ancestral species. Here we present an analytical treatment of the dynamics of the mean number of blocks over time, for continuous and discrete chromosomes, in finite and infinite populations. We find that we can describe the mean number of haplotype blocks as a universal function dependent on population size, the number of genomic elements, the length of the chromosome in Morgan, and the initial relative frequency of the ancestral species.

KEYWORDS junction; hybridization; haplotype blocks

Speciation through hybridization has long been recognized as a driving process in the formation of new species, in particular in plants (?). Only recently, have we begun to discover cases of speciation through hybridization in animals, including, but not limited to, butterflies (??), cichlid fishes (??), warblers (?), fruit flies (?) and sculpins (?). Speciation through hybridization, where the initial population of a species consists of hybrid individuals which are the result of a mating event between two related species, is thought to be a relatively rapid form of speciation, which can be completed within a few thousand generations

mating (?). In order to derive expected numbers of junctions, and variation in the number of junctions, Fisher had to assume that the length of the genetic blocks delineated by these junctions was exponentially distributed. Using simulations, ? showed that this assumption was inaccurate, and that although the block size distribution was close to exponential, large blocks tended to be overrepresented. Furthermore, Chapman and Thompson extended the theory of junctions towards populations growing in size at a constant rate, and towards subdividing populations (??).

and do not keep track of the identity of chromosomes. Each individual is diploid and chromosomes are inherited independently, which allows us to track haplotype blocks within only one chromosome pair, rather than all pairs simultaneously.

We start by formulating a recurrence equation based on the expected change, or the change in mean number of blocks per generation. Given a recombination site picked randomly across the length of the chromosome, the genomic material on either chromosome can either be identical, or different. If the genomic material is identical, no change in the number of blocks occurs. If the genomic material is different, a new block is formed (see figure ??). The probability of observing the same type of genomic material on both chromosomes is proportional to the frequency of genomic material of that type in the population. We denote the frequency of type P genomic material as p , and the frequency of genomic material of the other type Q as q , where $p = 1 - q$. The probability of having the same type of genomic material on both chromosomes at the recombination site is $p^2 + q^2$, in which case no change in the number of blocks is observed. With probability $2pq$ the type of genomic material on both chromosomes differs and an increase in number of blocks is observed. We obtain

$$\bar{n}_{t+1} = \bar{n}_t + 2pq. \quad (1)$$

Where \bar{n}_t is the average number of blocks at time t . The solution of equation ?? is given by

$$\bar{n}_t = n_0 + 2pqt. \quad (2)$$

The number of blocks increases linear in time. The probability of having a different type of genomic material, $2pq$ is possibly more familiar to the reader as the heterozygosity H , and we can write equation ?? in terms of heterozygosity, taking into account that the number of junctions J at time t is $J_t = n_t - 1$ and recover the previously obtained result (??):

$$E[J_t] = \sum_{j=0}^{t-1} H_j. \quad (3)$$

Where $E[J_t]$ is the expected number of junctions J at time t . Because the population size is infinite, the average heterozygosity does not change from H_0 , and $E[J_t] = H_0 t$. These results only apply for infinite population sizes. If the population size is constant over time and finite, and we allow for selfing, the average heterozygosity changes over time (?):

$$H_t = H_0 \prod_{i=0}^{t-1} \left(1 - \frac{1}{2N_i}\right). \quad (4)$$

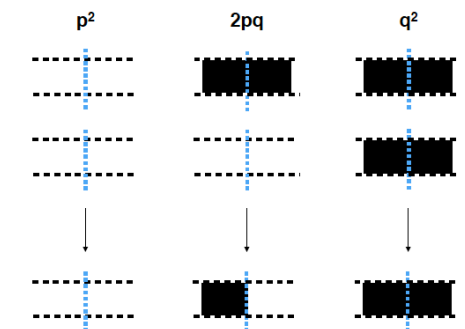


Figure 1 Change in number of haplotype blocks depending on the genomic match between blocks. Only a subset of a chromosome is shown, with in the middle the recombination site (blue dotted line), and on either end the color indicating the type of genomic material delineating the recombination site (type 1 = black, type 2 = white). With probability $p^2 + q^2$ no change in the number of blocks is observed. With probability $2pq$ we observe an increase in the number of blocks, where $p = 1 - q$ is the fraction of genomic material of type 1.

B. A finite number of recombination spots: a discrete chromosome

In the previous section we have assumed that crossover never occurs twice at the same spot. In reality, a chromosome can not be indefinitely divided into smaller parts. We therefore proceed to study the change in number of blocks in a chromosome consisting of L different chromosomal segments, where each segment represents a minimal genomic element that can not be broken down further, for instance a single nucleotide, a gene, a specific codon, or a genomic area delineated by two genetic markers, such as two Single Nucleotide Polymorphisms (SNPs). Considering a chromosome of L genomic segments, there are $L - 1$ possible crossover spots (junctions). Given that there are n blocks on the chromosome, there are $n - 1$ points on the chromosome where one block ends, and a new block begins (Fisher

Requirements

- Text editor (Notepad / TextEdit / BBEdit etc)
- Windows:
 - MikTeX
- Unix/Mac:
 - MacTeX
- Browser (all platforms):
 - Overleaf

Getting started

```
\documentclass{article}
```

```
\begin{document}
```

```
    Welcome to the LaTeX workshop.
```

```
\end{document}
```

- Available classes:
 - Article (what you will typically use for scientific articles, short texts etc)
 - Book (suitable for thesis, has support for chapters, middle margins, front matter etc)
 - Beamer (used for presentations)
 - Letter (for writing letters)
 - And many more: <https://www.ctan.org/topic/class>

Bold and italic

```
\documentclass{article}
```

```
\begin{document}
```

```
Welcome to the LaTeX workshop. \\
```

```
We can \textit{emphasize} \textbf{important} words  
easily.
```

```
\end{document}
```

Sections

```
\section{Introduction}
```

```
\section{Math}
```

```
\section{Objects}
```

```
\subsection{Figures}
```

```
\subsection{Tables}
```

```
\section{Referencing}
```


Adding equations

Inline equation:

The frequency is given by $1 - p$.

Numbered equation:

And thus we find:

```
\begin{equation}
```

$$(1 - p) + 2pq + (1 - q) = 1$$

```
\end{equation}
```

Adding a figure

```
\usepackage{graphicx}
```

```
\begin{figure}
```

```
    \centering
```

```
    \includegraphics{gopher_438742864.jpeg}
```

```
    \caption{This is a gopher}
```

```
    \label{fig:gopher}
```

```
\end{figure}
```

Figure tags

```
\begin{figure} [ 'placement specifier' ]  
..  
figure code  
\\end{figure}
```

- Placement specifiers:
 - **h**: place the figure *here*
 - **t**: place the figure at the *top* of the page
 - **b**: place the figure at the *bottom* of the page
 - **p**: place the figure on it's own *page*

Adding a table

```
\begin{table}[h]
  \begin{center}
    \begin{tabular}{| c | c | c |}
      \hline
      education & experienced & novice \\
      \hline\hline
      PhD & 5 & 10 \\
      \hline
      post-Doc & 3 & 1 \\
      \hline
    \end{tabular}
    \caption{Data from the Workshop}
    \label{tab:workshop_data}
  \end{center}
\end{table}
```

Column types

- **c**: centered
- **r**: right-aligned
- **l**: left-aligned

Cross referencing

You can add labels to anything (sections, figures, tables, equations).

Then, using `\ref` you can refer to that label.

In section `\ref{sec:tables}`, we showed that many of the participants are not experienced with `\LaTeX` yet (See Table `\ref{tab:workshop_data}`). However, Gopher (see Figure `\ref{fig:gopher_figure}`) has made sure that this workshop will remedy that.

Citing + Bibliography

Citations organised in a 'BiBTeX file':

```
\cite{huntly1988pocket}
```

```
\section{Bibliography}
```

```
\bibliographystyle{apalike}
```

```
\bibliography{references}
```

```
@article{huntly1988pocket,  
  title={Pocket gophers in  
ecosystems: patterns and mechanisms},  
  author={Huntly, Nancy and Inouye,  
Richard},  
  journal={BioScience},  
  volume={38},  
  number={11},  
  pages={786--793},  
  year={1988},  
  publisher={JSTOR}  
}
```

Concluding remarks

- LaTeX may take some time to get used to, Google is your friend!
- Useful for:
 - Large manuscripts
 - Equation / Math heavy manuscripts
 - Making beautiful tables
 - Having a text with beautiful typesetting

Pimping your ride: modifying fonts

- Increase font size:
 - `\documentclass[12pt]{article}`
 - By default, this is 10pt, you can choose from 10, 11 and 12pt
 - If you want more different sizes, you can use the package 'ext sizes', and pick from: 9, 10, 11, 12, 14, 17 and 20pt.
- If you want to change font size locally, you can use:
- `\Huge`, `\huge`, `\Large`, `\large`, `\small`, `\tiny`:
 - `{\Huge this text is very large}`
 - `{\tiny this text is very small}`
- Notice the use of bracketing to indicate the scope of you font size!

Pimping your ride: margins

- Margins can be modified using the 'Geometry' package:

```
\usepackage{geometry}  
\geometry{a4paper, left=20mm, top=20mm}
```

This sets the margins on the left of the text to 20mm, and between the text and the top of the page to 20mm. Other options are bottom and right (but these are kept default here).

- Please note that margins might interact with some other packages! For instance, the 'landscape' package, forces margins to be the default margins.

Multiple columns

- In the entire document:

```
\documentclass[10pt, twocolumn]{article}
```

- In a section:

```
\usepackage{multicol}
```

```
\begin{multicols}{2}
```

Focal text

```
\end{multicols}{2}
```

Landscape section

- Sometimes, your table is simply too wide. Can we add a section in landscape?

```
\usepackage{lscape}
```

```
\begin{landscape}
```

```
% here goes your table content
```

```
\end{landscape}
```

- Please note that this might interact / interfere with margins!

Journal templates

- Instead of modifying everything by hand, we can also pick from journal templates in Overleaf!
- <https://www.overleaf.com/latex/templates>