Title of Thesis
Title of Thesis Line 2
Title of Thesis Line3

Author of Thesis

Doctor of Philosophy

Department

McGill University

Montreal,Quebec

2004-12-01

DEDICATION

This document is dedicated to the graduate students of the McGill University.

ACKNOWLEDGEMENTS

Acknowledgments, if included, must be written in complete sentences. Do not use direct address. For example, instead of Thanks, Mom and Dad!, you should say I thank my parents.

ABSTRACT

Abstract in English and French are required. The text of the abstract in English begins here.

ABRÉGÉ

The text of the abstract in French begins here.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Introduction

How does one's DNA influence their risk of getting a disease? Contrary to popular belief, your future health is not "hard wired" in your DNA. Only in a few diseases, referred as "Mendelian diseases", are there well known, almost certain, links between genetic mutations and disease susceptibility. For the majority of what are known as "complex traits", such as cancer or diabetes, genomic predisposition is subtle and, so far, not fully understood.

With the rapid decrease in the cost of DNA sequencing, the complete genome sequence of large cohorts of individuals can now be routinely obtained. This wealth of sequencing information is expected to ease the identification of genetic variations linked to complex traits. In this work, I investigate the analysis of genomic data in relation to complex diseases, which offers a number of important computational and statistical challenges. We tackle several steps necessary for the analysis of sequencing data and the identification of links to disease. Each step, which corresponds to a chapter in my thesis, is characterized by very different problems that need to be addressed.

i) The first step is to analyze large amounts of information generated by DNA sequencers to obtain a set of "genomic variants" present n each each individual. To address these big data processing problems, Chapter **??** shows how

1

we designed a programming language (BigDataScript [**?**]), that simplifies the creation robust, scalable data pipelines.

ii) Once genomic variants are obtained, we need to prioritize and filter them to discern which variants should be considered "important" and which ones are likely to be less relevant. We created the SnpEff & SnpSift [**?**, **?**] packages that, using optimized algorithms, solve several annotation problems: a) standardizing the annotation process, b) calculating putative genetic effects, c) estimating genetic impact, d) adding several sources of genetic information, and e) facilitating variant filtering.

iii) Finally, we address the problem of finding associations between interacting genetic loci and disease. One of the main problems in GWAS, known as "missing heritability", is that most of the phenotypic variance attributed to genetic causes remains unexplained. Since interacting genetic loci (epistasis) have been pointed out as one of the possible causes of missing heritability, finding links between such interactions and disease has great significance in the field. We propose a methodology to increase the statistical power of this type of approaches by combining population-level genetic information with evolutionary information.

In a nutshell, this thesis addresses computational, analytical, algorithmic and methodological problems of transforming raw sequencing data into biological insight in the aetiology of complex disease. In the rest of this introduction we give the background that provides motivation for our research. The idea of the McGill ETD class (mcgilletdclass) is to input the standard LaTeX 2_ε report class and make changes

as necessary to meet the requirements of the McGill EThesis Project. The following packages, in addition to those included with the base distribution of $\text{LaTeX}\,2_\varepsilon$, must exist on your system; the dates in parentheses give the dates of the oldest version this class was tested with. Since they are all standard packages, they are readily available at any CTAN (Comprehensive TEX Archive Network) site.

- setspace (2000/12/01)
- ulem (2000/05/26)

  You can use `\uline{text}` to underline <u>text</u>. The documentation of the ulem package provides more information.

- sectsty (1999/04/12)

  Since underlining in section headings is rather nontrivial, the sectsty package was used to manipulate the formatting of sectional headings.

- ragged2e (1999/06/08)

  This package adds support to handle ragged right justification appropriately.

- everysel (1999/06/08)

  Required by the ragged2e package.

1.2 General Guidelines

- Please use commands provided by mcgilletdclass, otherwise your thesis may not be able to pass the validation check.
- If you run into problems using this stylesheet, please contact us as soon as possible.
- Files you must submit: *.tex, *.dvi and *.bib files.

CHAPTER 2
Usage

An example of how to use the mcgilletdclass document class is provided in mcgilletdsample.tex. Details of how to compile the example are provided in a later section. Use this class file the same way as the report class, by putting `\documentclass{mcgilletdclass}` at the beginning of the LaTeX 2_ε file. There are only a few options available with this document class, which will be described in the following sections. Due to the requirements of the McGill University, this document class does not support **two-sided printing or two-column page layout** .

CHAPTER 3
The document class options

3.1 Font Size Options: 10pt, 11pt or 12pt

These are the font sizes that are available in the standard LaTeX 2_ε report document class. The 12pt font size is the preferred option for the final typeset version of the document. Obviously, only one of these options should be used in the `\documentclass` command. The 12pt option is selected by default if none of these options are specified in the `\documentclass` command.

CHAPTER 4
Title Page

The document preamble is the part that occurs before the `\begin{document}` statement. Throughout the document, information about the author, the title, etc., is **required** . The following commands are used in the document preamble to define all of the required text strings that are used to personalize the document.

- `\SetTitle{text}` :This is the title as it appears on the title page. The title must be entered using all upper-case letters. Line-breaks can be entered in the title by using the `\\` command.

- `\SetAuthor{text}` :The author's name, using capital letters where appropriate.

- `\SetDegreeType{text}` :Something like "Master of Science" or "Doctor of Philosophy".

- `\SetDepartment{text}` :The department or faculty

- `\SetUniversity{text}` :McGill University

- `\SetUniversityAddr{text}` :Montreal, Quebec

- `\SetThesisDate{text}` :Date of your thesis

- `\SetRequirements{text}` :Text of requirements statements

- `\SetCopyright{text}` :Copyright Statements

CHAPTER 5
The Thesis FrontSection

In this section, we present the commands used to add all pages prior to the first chapter of the thesis. Some of these commands are required while some of them are optional.

5.1   Title page(required)

The command `\maketitle` formats the title page.

5.2   Dedication

The optional dedication can be inserted by using the command `\SetDedicationName{text},\SetD` and `\Dedication`, where text is the contents of the dedication.

5.3   Acknowledgments

In order to format the optional acknowledgments, use the command `\SetAcknowledgeName{text}` and `\Acknowledge{text}`. The title "ACKNOWLEDGMENTS" is automatically added to the table of contents.

5.4   Table of Contents, List of Tables and Figures(required)

These lists are generated with the commands `\tableofcontents`, `\listoftables and \listoffi` The titles of these lists may be changed by renewing the `\TOCHeading{text}`, `\LOTHeading{text}`, and the `\LOFHeadling{text}` macros, respectively.

5.5   Abstract(required)

This command, which actually is an environment, sets up the required text for the abstract. The proper use is : `\SetAbstractEnName{text}` for the heading of the

abstract and `\SetAbstractEnText{text}` for the text of the abstract. The command `\AbstractEn` is used to incorporate the heading of the abstract and the abstract text into your document. `\SetAbstractFrName{text}`,`\SetAbstracFrText{text}` and `\AbstractFr` are the equivalent commands for the French abstract, and `\SetAbstractOtName{text}`, `\SetAbstracOtText{text}` and `\AbstractOt` are for an abstract in another language. And abstract in English and French are required by the University.

CHAPTER 6
The Thesis "Mainmatter"

In this section, we describe the commands used to format the main chapters of the thesis. Please refer to the sample thesis file mcgilletdsample.tex for examples of how these commands are used.

6.1 Chapter Titles

Use the \chapter command to start a new chapter in the document. If the boolean SetDSpace is true, then the chapter will be typeset using double-spacing. In general, the class file sets the SetDSpace boolean appropriately so that the selection of singlespacing or double-spacing is transparent to the user. Note that the chapter title may extend over more than one line by using the following form: \chapter[Title Line 1 \protect\newline Title Line 2]Title Line 1 \\ Title Line 2

6.2 Section Headings

Use the \section command for first-level subheadings. Use the \subsection command for second-level subheadings. Since it may be difficult to distinguish third-level subheadings from second-level subheadings, it is suggested that the \paragraph command be used for all third-level subheadings. Note that the text of the heading may extend over more than one line. In this case, use the same command as shown for the \chapter command above.

6.3   Tables and Figures

Use the `\begin{table}` `and` `\end{table}` to contain a new table.
Use the `\begin{figure}` `and` `\end{figure}` to include a new figure. Captions must
be placed above tables and below figures. `\TableCaption{text}` is for tables and
`\FigureCaption{text}` is for figures. Please note that captions are required for
both figures and tables, and must be supplied on a strictly one-to-one basis.

`\TableCaptionOpt{text1}{text2}` is another command for the similar use.
Here, text1 will appear in the list of tables; text2 is the actual caption of a table.
Similarly, `\FigureCaptionOpt{text1}{text2}` is for figures.

CHAPTER 7
The Thesis "Backmatter"

The following commands are used to control the appearance and content of appendices and other thesis end-matter. You might have to tailor some of these commands to suit your requirements. Use the `\ETDAppendix{Appendix}{text}` and `\ETDAppendix{Appendix}{text}` commands to add an appendix; you may have more than one appendix if necessary.

7.1 List of References

We only support the BibTex tool. Use the `\bibHeading{text}` command to set the heading of your references. Use the `\bibliography{bibtexdatabase}` to incorporate your BibTex list. Use the `\bibliographystyle{text}` to set the citation style of your bibliography.

Here is a simple example to demonstrate the use of BibTex.

The following text and latex commands (in conjunction with the mgilletd.bib file included in the mcgilletd package):

These rather startling results were explained by subsequent research `~\cite{Ammons:001}` and were later separately confirmed `~\cite{Fishman:001}`.

produce:

These rather startling results were explained by subsequent research [?] and were later separately confirmed [?].

The references themselves are inserted in the document at the point of the `\bibliography{mcgilletd}` command (see the end of this document).

CHAPTER 8
Special Commands

The commands described in this section are used to aid in the commenting of the source file, and also describe support for special formatting. These commands are optional, in that they are not required in order to format the document.

8.1 Automatically Generated Indices

Several commands have been included in this document class file to facilitate the creation of automatically generated lists with the MakeIndex program. Use this command in the document preamble to enable the creation of an index file. If only one index is to be created, then just use the `\makeindex` command. However, if there is to be more than one index, or it is desirable to name the index file something besides the default, use the following command `\makeindex[filename]` A unique filename for each index must be assigned, so a `\makeindex[]` command should exist for each index, in the document preamble. Issue this command at the point in the document where the index is to be typeset. Typically, this will be after the lists of figures and tables (if included). The command creates a new chapter and sets up the initial page formatting for the list to be generated. The form of the `\printindex` command directly corresponds to the form of the `\makeindex` command used in the document. This is the command that is used to actually define what is going to be added to the index list. As with the other indexing commands

there are two forms that can be used, depending on how the `\makeindex` command was issued. For generating only one index file, use the command shown below `\index`text to add to index If there are more than one index files, or if a custom filename was used for the index, then the following form of the command must be used: `\index[filename]{text to add to index}` Of course, for each index file that is created, a run of the MakeIndex program will be required.

8.2 Changing the Code . . .

It might be necessary to modify some of the commands defined in the documentclass. However, you do NOT have to (and SHOULD NOT!) edit the mcgilletdclass.cls file directly. Instead, place any commands you wish to create or modify in a file called mcgilletdclass.cfg and put it in the same directory as your main thesis file. The mcgilletdclass documentclass will automatically load this file. Please note that required commands must not be changed if your thesis is to be correctly converted to XML for the McGill EThesis Project.

8.3 An Example Thesis File

In order to typeset the example file, the following files should exist: mcgilletdclass.cls, mcgilletdsample.tex, mcgilletd.bib, chick.eps, and the additional packages described at the beginning of this document. (See the documentation for your local installation of LaTeX 2_ε for information on how to install local customizations.)

Use the following commands to typeset the mcgilletdsample file:

```
latex mcgilletdsample
bibtex mcgilletdsample
latex mcgilletdsample
```

```
latex mcgilletdsample
```

8.4 Special Cases

An objective of the McGill Electronic Thesis Project is to create a means to ensure the long-term preservation and long-term accessibility of theses created at McGill University. The submitted theses will be archived as XML, which is a subset of SGML, a markup language for the interchange of structured data. The basis for SGML/XML is the description of the contents, not the layout. Therefore, the mcgilletdclass documentclass adds sematic support to text layout commands. Some of them are described as below:

- Emphasis ProperName

  i `\BProperName` is to set text to boldface; the text will be treated as a proper name.

  ii `\BIProperName` is to set text to boldface and italic; the text will be treated as a proper name.

  iii `\BIUProperName` is to set text to boldface, italic and underline; the text will be treated as a proper name.

  iv `\BUProperName` is to set text to boldface and underline; the text will be treated as a proper name.

  v `\IUProperName` is to set text to italic and underline; the text will be treated as a proper name.

  vi `\UProperName` is to set text to underline; the text will be treated as a proper name.

vii `\IProperName` is to set text to italic; the text will be treated as a proper name.

- Emphasis Technical Term

    i `\BTechnicalTerm` is to set text to boldface; the text will be treated as a technical term.

    ii `\BITechnicalTerm` is to set text to boldface and italic; the text will be treated as a technical term.

    iii `\BIUTechnicalTerm` is to set text to boldface, italic and underline; the text will be a treated as technical term.

    iv `\BUTechnicalTerm` is to set text to boldface and underline; the text will be treated as a technical term.

    v `\IUTechnicalTerm` is to set text to italic and underline; the text will be treated as a technical term.

    vi `\UTechnicalTerm` is to set text to underline; the text will be treated as a technical term.

    vii `\ITechnicalTerm` is to set text to italic; the text will be treated as a technical term.

- Emphasis Title

    i `\BTitle` is to set text to boldface; the text will be treated as a title.

    ii `\BITitle` is to set text to boldface and italic; the text will be treated as a title.

    iii `\BIUTitle` is to set text to boldface, italic and underline; the text will be a treated as title.

iv \BUTitle is to set text to boldface and underline; the text will be treated as a title.

v \IUTitle is to set text to italic and underline; the text will be treated as a title.

vi \UTitle is to set text to underline; the text will be treated as a title.

vii \ITitle is to set text to italic; the text will be treated as a title.

- Emphasis ForeignWord

i \BForeignWord is to set text to boldface; the text will be treated as a Foreign Word.

ii \BIForeignWord is to set text to boldface and italic; the text will be treated as a Foreign Word.

iii \BIUForeignWord is to set text to boldface, italic and underline; the text will be a treated as Foreign Word.

iv \BUForeignWord is to set text to boldface and underline; the text will be treated as a Foreign Word.

v \IUForeignWord is to set text to italic and underline; the text will be treated as a Foreign Word.

vi \UForeignWord is to set text to underline; the text will be treated as a Foreign Word.

vii \IForeignWord is to set text to italic; the text will be treated as a Foreign Word.

- Visual Emphasis

i \BVisualEmphasis is to set text to boldface; the text will be treated as Visual Emphasis.

ii \BIVisualEmphasis is to set text to boldface and italic; the text will be treated as Visual Emphasis.

iii \BIUVisualEmphasis is to set text to boldface, italic and underline; the text will be treated as Visual Emphasis.

iv \BUVisualEmphasis is to set text to boldface and underline; the text will be treated as Visual Emphasis.

v \IUVisualEmphasis is to set text to italic and underline; the text will be treated as Visual Emphasis.

vi \UVisualEmphasis is to set text to underline; the text will be treated as Visual Emphasis.

vii \IVisualEmphasis is to set text to italic; the text will be treated as Visual Emphasis.

- Quotation

  i Use \SingleInlineQuote command to do 'inline quote'.

  ii Use \DoubleInlineQuote command to do "inline quote".

  iii Use \begin{quote} and \end{quote} for block quote

- Verse

  i Use \verse command for verse.

- Lists

  i Use \begin{BulletList} and \end{BulletList} commands to create bulleted lists.

ii Use \begin{RomanList} and \end{RomanList} commands to create upper case Roman numbered lists.

iii Use \begin{romanList} and \end{romanList} commands to create lower case roman numbered lists.

iv Use \begin{AlphList} and \end{AlphList} commands to create upper case Alphabetical lists.

v Use \begin{alphList} and \end{alphList} commands to create lower case alphabetical lists.

vi Use \begin{arabicList} and \end{arabicList} commands to create arabic numbered lists.

- Formula

i Use \begin{MathFormula} and \end{MathFormula} commands to embed normal math formula as per the following example:

\begin{MathFormula}

\begin{math}

...

\end{math}

\end{MathFormula}

Example:

$$\bar{u} \;=\; \int_I^E g(x)dx$$

(8.1)

ii Use \begin{ChemFormula} and \end{ChemFormula} commands for chemical formulae, similar to the above example.

\begin{ChemFormula}

...

\end{ChemFormula}

8.5   Testing figures and tables.

Table 8–1: The First Table

First	Second	Third	Fourth
12	26	12	33
17	93	88	3
12	26	12	33
12	26	12	33
12	26	12	33
17	93	88	3
12	26	12	33
12	26	12	33
12	26	12	33
17	93	88	3
12	26	12	33
12	26	12	33

Figure 8–1: My Little Chick

Figure 8–2: A Resized Figure

CHAPTER 9
TESTING FOOTNOTES

We will now work on testing footnotes.[1] And now let's add some spurious text so that we can have a second[2] footnote.

---

[1] This is the first footnote.

[2] This is the second footnote. To test if the footnote is typeset single-spaced, we add meaningless text until we can see a second line.

CHAPTER 10
Generating an Index

10.1　Genereating an Index

If you want to generate an index, you will need the command `\makeindex[index]` in the preamble. Here, index can be any word named by users.

The command `\index[index]{entry}` is used to index entry at that point in the document. For example, the following code:

`\makeindex[myindex1]` ...

... Eigenvectors`\index[myindex1]{eigenvector}` ...

will produce output and place the entry 'eigenvector' in the myindex1.idx file with the associated page number.

The mcgill etd class redefines the command `\printindex` which should be placed at the point in the document where you want your index to appear.And `\printindex` command can be used as following:

`\printindex[myindex1]{Index - appears in the table of contents}`

`{Index - will be the title of your index.}`

`{Some textual information you want to add the head of your index.}`

Two index formats are defined. One is with leading dots, another is without dots. If you are also using BIBTEX and index without leading dots, you will need to do:

latex filename.tex

bibtex filename

makeindex -s mcgillpage.ist myindex1

latex filename.tex

latex filename.tex

10.2 Generating a list of abbreviations

Generating a list of abbreviations is similar to generating an index. Usually, items of an index follow by a page number. However, lists of abbreviations do not include page number at all. In this case, you have to use the following commands to generate a list of abbreviation:

`\makeindex[ListOfAbbr]`

`\index[ListOfAbbr]{WWW@World Wide Web}`

...

`\printindex[ListOfAbbr]{List of Abbreviation}{List of Abbreviation}{}`

You need to use makeindex to generate the list as below:

makeindex -s mcgillnopage.ist ListOfAbbr

CHAPTER 11
Supporting TexShade

If you want to use TexShade, you will need to add an option "TexShade" to the `\documentclass[optionlist]{mcgilletdclass}`. And it looks like the following: `\documentclass[TexShade, 12pt, letterpaper]{mcgilletdclass}`. Then you should add the TexShade package into your thesis: `\usepackage{TexShade}`. Here is an example:

`\begin{texshade}{alng.fasta}`

`\noblockskip`

`\hideconsensus`

`\end{texshade}`

And you will get the following result:

alng.fasta

Appendix A

Here is the text of an Appendix.

Appendix B

Here is the text of a second, additional Appendix