TARGETED NEXT GENERATION SEQUENCING PANEL FOR CLINICAL DIAGNOSTICS OF SOLID TUMOURS

by

Shyong Quin Yap

B.Sc. (Hons), Trent University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Experimental Medicine Program)

The University of British Columbia

(Vancouver)

January 2017

Abstract

This document provides brief instructions for using the `ubcdiss` class to write a UBC-conformant dissertation in LATEX. This document is itself written using the `ubcdiss` class and is intended to serve as an example of writing a dissertation in LATEX. This document has embedded Unique Resource Locators (URLS) and is intended to be viewed using a computer-based Portable Document Format (PDF) reader.

Note: Abstracts should generally try to avoid using acronyms.

Note: at University of British Columbia (UBC), both the Graduate and Postdoctoral Studies (GPS) Ph.D. defence programme and the Library's online submission system restricts abstracts to 350 words.

Preface

At UBC, a preface may be required. Be sure to check the GPS guidelines as they may have specific content to be included.

Table of Contents

List of Tables

List of Figures

Glossary

This glossary uses the handy `acroynym` package to automatically maintain the glossary. It uses the package's `printonlyused` option to include only those acronyms explicitly referenced in the LaTeX source.

CTAN The Common TeX Archive Network

DOI Document Object Identifier (see http://doi.org)

GPS Graduate and Postdoctoral Studies

PDF Portable Document Format

RCS Revision control system, a software tool for tracking changes to a set of files

URL Unique Resource Locator, used to describe a means for obtaining some resource on the world wide web

Acknowledgments

Although this thesis only bears one name, its completion would be impossible without the contribution of many individuals. First and foremost, I would like to express my sincere gratitute to my supervisor, Dr. Aly Karsan, for the opportunity to work in his lab which consists of diverse talents as well as his patience, guidance, and extensive knowledge in clinical informatics.

I would also like to thank my labmates from the bioinformatics team, Kieran and Rod, and members from the Centre of Clinical Genomics, Liz and Jill, for their insightful comments and help throughout my time in the lab. I would like to extend my gratitude to my supervisory committee members, Dr. Ryan Morin and Dr. Martin Hirst, for their knowledgeable feedback and continuous effort in asking me difficult questions which motivated me to widen my research from various perspectives.

My sincere thanks also goes to my friends who supported me and lifted my spirits through the tough times. Last but not least, I would like to thank my family: mom and dad, for always encouraging my interest in science and listening to my endless science talks and my sisters, for believing in my ability even when I doubt myself. I am a lucky person.

Chapter 1

Introduction

> *If I have seen farther it is by standing on the shoulders of Giants.* — Sir Isaac Newton
> (1855)

This document provides a quick set of instructions for using the `ubcdiss` class to write a dissertation in LaTeX. Unfortunately this document cannot provide an introduction to using LaTeX. The classic reference for learning LaTeX is Lamport's book [3]. There are also many freely-available tutorials online; Andy Roberts' online LaTeX tutorials[1] seems to be excellent. The source code for this docment, however, is intended to serve as an example for creating a LaTeX version of your dissertation.

We start by discussing organizational issues, such as splitting your dissertation into multiple files, in Section 1.1. We then cover the ease of managing cross-references in LaTeX in Section 1.2. We cover managing and using bibliographies with BibTeX in Section 1.3. We briefly describe typesetting attractive tables in Section 1.4. We briefly describe including external figures in Section 1.5, and using special characters and symbols in Section 1.6. As it is often useful to track different versions of your dissertation, we discuss revision control further in Section 1.8. We conclude with pointers to additional sources of information in Section 1.10.

1.1 Suggested Thesis Organization

The UBC Graduate and Postdoctoral Studies (GPS) specifies a particular arrangement of the components forming a thesis.[2] This template reflects that arrangement.

In terms of writing your thesis, the recommended best practice for organizing large documents in LaTeX is to place each chapter in a separate file. These chapters are then included from the main file through the use of `\include{file}`. A thesis might be described as six files such as `intro.tex`, `relwork.tex`, `model.tex`, `eval.tex`, `discuss.tex`, and `concl.tex`.

---

[1] http://www.andy-roberts.net/misc/latex/

[2] See http://www.grad.ubc.ca/current-students/dissertation-thesis-preparation/order-components

We also encourage you to use macros for separating how something will be typeset (e.g., bold, or italics) from the meaning of that something. For example, if you look at `intro.tex`, you will see repeated uses of a macro `\file{}` to indicate file names. The `\file{}` macro is defined in the file `macros.tex`. The consistent use of `\file{}` throughout the text not only indicates that the argument to the macro represents a file (providing meaning or semantics), but also allows easily changing how file names are typeset simply by changing the definition of the `\file{}` macro. `macros.tex` contains other useful macros for properly typesetting things like the proper uses of the latinate *exempli gratiā* and *id est* (i.e., `\eg` and `\ie`), web references with a footnoted URL (`\webref{url}{text}`), as well as definitions specific to this documentation (`\latexpackage{}`).

1.2   Making Cross-References

LATEX make managing cross-references easy, and the `hyperref` package's `\autoref{}` command[3] makes it easier still.

A thing to be cross-referenced, such as a section, figure, or equation, is *labelled* using a unique, user-provided identifier, defined using the `\label{}` command. The thing is referenced elsewhere using the `\autoref{}` command. For example, this section was defined using:

```
\section{Making Cross-References}
\label{sec:CrossReferences}
```

References to this section are made as follows:

```
We then cover the ease of managing cross-references in \LaTeX\
in \autoref{sec:CrossReferences}.
```

`\autoref{}` takes care of determining the *type* of the thing being referenced, so the example above is rendered as

> We then cover the ease of managing cross-references in LATEX in Section 1.2.

The label is any simple sequence of characters, numbers, digits, and some punctuation marks such as ":" and "–"; there should be no spaces. Try to use a consistent key format: this simplifies remembering how to make references. This document uses a prefix to indicate the type of the thing being referenced, such as `sec` for sections, `fig` for figures, `tbl` for tables, and `eqn` for equations.

For details on defining the text used to describe the type of *thing*, search `diss.tex` and the `hyperref` documentation for `autorefname`.

1.3   Managing Bibliographies with BibTEX

One of the primary benefits of using LATEX is its companion program, BibTEX, for managing bibliographies and citations. Managing bibliographies has three parts: (i) describing references, (ii) cit-

---

[3]The `hyperref` package is included by default in this template.

ing references, and (iii) formatting cited references.

1.3.1 Describing References

BibTEX defines a standard format for recording details about a reference. These references are recorded in a file with a `.bib` extension. BibTEX supports a broad range of references, such as books, articles, items in a conference proceedings, chapters, technical reports, manuals, dissertations, and unpublished manuscripts. A reference may include attributes such as the authors, the title, the page numbers, the Document Object Identifier (DOI), or a Unique Resource Locator (URL). A reference can also be augmented with personal attributes, such as a rating, notes, or keywords.

Each reference must be described by a unique *key*.[4] A key is a simple sequence of characters, numbers, digits, and some punctuation marks such as ":" and "–"; there should be no spaces. A consistent key format simiplifies remembering how to make references. For example:

$$\boxed{\textit{last-name}}\text{–}\boxed{\textit{year}}\text{–}\boxed{\textit{contracted-title}}$$

where *last-name* represents the last name for the first author, and *contracted-title* is some meaningful contraction of the title. Then Kiczales et al.'s seminal article on aspect-oriented programming [2] (published in 1997) might be given the key `kiczales-1997-aop`.

An example of a BibTEX `.bib` file is included as `biblio.bib`. A description of the format a `.bib` file is beyond the scope of this document. We instead encourage you to use one of the several reference managers that support the BibTEX format such as JabRef[5] (multiple platforms) or BibDesk[6] (MacOS X only). These front ends are similar to reference manages such as EndNote or RefWorks.

1.3.2 Citing References

Having described some references, we then need to cite them. We do this using a form of the `\cite` command. For example:

```
\citet{kiczales-1997-aop} present examples of crosscutting
from programs written in several languages.
```

When processed, the `\citet` will cause the paper's authors and a standardized reference to the paper to be inserted in the document, and will also include a formatted citation for the paper in the bibliography. For example:

> Kiczales et al. [2] present examples of crosscutting from programs written in several languages.

---

[4]Note that the citation keys are different from the reference identifiers as described in Section 1.2.

[5]http://jabref.sourceforge.net

[6]http://bibdesk.sourceforge.net

Table 1.1: Available `cite` variants; the exact citation style depends on whether the bibliography style is numeric or author-year.

Variant	Result
`\cite`	Parenthetical citation (e.g., "[2]" or "(Kiczales et al. 1997)")
`\citet`	Textual citation: includes author (e.g., "Kiczales et al. [2]" or or "Kiczales et al. (1997)")
`\citet*`	Textual citation with unabbreviated author list
`\citealt`	Like `\citet` but without parentheses
`\citep`	Parenthetical citation (e.g., "[2]" or "(Kiczales et al. 1997)")
`\citep*`	Parenthetical citation with unabbreviated author list
`\citealp`	Like `\citep` but without parentheses
`\citeauthor`	Author only (e.g., "Kiczales et al.")
`\citeauthor*`	Unabbreviated authors list (e.g., "Kiczales, Lamping, Mendhekar, Maeda, Lopes, Loingtier, and Irwin")
`\citeyear`	Year of citation (e.g., "1997")

There are several forms of the `\cite` command (provided by the `natbib` package), as demonstrated in Table 1.1. Note that the form of the citation (numeric or author-year) depends on the bibliography style (described in the next section). The `\citet` variant is used when the author names form an object in the sentence, whereas the `\citep` variant is used for parenthetic references, more like an end-note.

1.3.3   Formatting Cited References

BibTeX separates the citing of a reference from how the cited reference is formatted for a bibliography, specified with the `\bibliographystyle` command. There are many varieties, such as `plainnat`, `abbrvnat`, `unsrtnat`, and `vancouver`. This document was formatted with `abbrvnat`. Look through your TeX distribution for `.bst` files. Note that use of some `.bst` files do not emit all the information necessary to properly use `\citet{}`, `\citep{}`, `\citeyear{}`, and `\citeauthor{}`.

There are also packages available to place citations on a per-chapter basis (`bibunits`), as footnotes (`footbib`), and inline (`bibentry`). Those who wish to exert maximum control over their bibliography style should see the amazing `custom-bib` package.

1.4   Typesetting Tables

Lamport [3] made one grievous mistake in LaTeX: his suggested manner for typesetting tables produces typographic abominations. These suggestions have unfortunately been replicated in most

LATEX Rocks!

Figure 1.1: Proof of LATEX's amazing abilities

LATEX tutorials. These abominations are easily avoided simply by ignoring his examples illustrating the use of horizontal and vertical rules (specifically the use of `\hline` and `|`) and using the `booktabs` package instead.

The `booktabs` package helps produce tables in the form used by most professionally-edited journals through the use of three new types of dividing lines, or *rules*. Tables 1.1 and 1.2 are two examples of tables typeset with the `booktabs` package. The `booktabs` package provides three new commands for producing rules: `\toprule` for the rule to appear at the top of the table, `\midrule` for the middle rule following the table header, and `\bottomrule` for the bottom-most at the end of the table. These rules differ by their weight (thickness) and the spacing before and after. A table is typeset in the following manner:

```
\begin{table}
\caption{The caption for the table}
\label{tbl:label}
\centering
\begin{tabular}{cc}
\toprule
Header & Elements \\
\midrule
Row 1 & Row 1 \\
Row 2 & Row 2 \\
% ... and on and on ...
Row N & Row N \\
\bottomrule
\end{tabular}
\end{table}
```

See the `booktabs` documentation for advice in dealing with special cases, such as subheading rules, introducing extra space for divisions, and interior rules.

1.5 Figures, Graphics, and Special Characters

Most LATEX beginners find figures to be one of the more challenging topics. In LATEX, a figure is a *floating element*, to be placed where it best fits. The user is not expected to concern him/herself with the placement of the figure. The figure should instead be labelled, and where the figure is used, the text should use `\autoref` to reference the figure's label. Figure 1.1 is an example of a figure. A figure is generally included as follows:

```
\begin{figure}
\centering
\includegraphics[width=3in]{file}
```

```
\caption{A useful caption}
\label{fig:fig-label}    % label should change
\end{figure}
```

There are three items of note:

1. External files are included using the `\includegraphics` command. This command is defined by the `graphicx` package and can often natively import graphics from a variety of formats. The set of formats supported depends on your TEX command processor. Both `pdflatex` and `xelatex`, for example, can import GIF, JPG, and PDF. The plain version of `latex` only supports EPS files.

2. The `\caption` provides a caption to the figure. This caption is normally listed in the List of Figures; you can provide an alternative caption for the LoF by providing an optional argument to the `\caption` like so:

   ```
   \caption[nice shortened caption for LoF]{%
       longer detailed caption used for the figure}
   ```

   GPS generally prefers shortened single-line captions in the LoF: multiple-line captions are a bit unwieldy.

3. The `\label` command provides for associating a unique, user-defined, and descriptive identifier to the figure. The figure can be can be referenced elsewhere in the text with this identifier as described in Section 1.2.

See Keith Reckdahls excellent guide for more details, *Using imported graphics in LaTeX2e*[7].

1.6   Special Characters and Symbols

LATEX appropriates many common symbols for its own purposes, with some used for commands (i.e., `\ { } & %`) and mathematics (i.e., `$ ^ _`), and others are automagically transformed into typographically-preferred forms (i.e., – '') or to completely different forms (i.e., <>). Table 1.2 presents a list of common symbols and their corresponding LATEX commands. A much more comprehensive list of symbols and accented characters is available at: http://www.ctan.org/tex-archive/info/symbols/comprehensive/

1.7   Changing Page Widths and Heights

The `ubcdiss` class is based on the standard LATEX `book` class that selects a line-width to carry approximately 66 characters per line. This character density is claimed to have a pleasing appearance and also supports more rapid reading [1]. I would recommend that you not change the line-widths!

---

[7]http://www.ctan.org/tex-archive/info/epslatex.pdf

Table 1.2: Useful LaTeX symbols

LaTeX	Result	LaTeX	Result
`\texttrademark`	™	`\&`	&
`\textcopyright`	©	`\{ \}`	{ }
`\textregistered`	®	`\%`	%
`\textsection`	§	`\verb!~!`	~
`\textdagger`	†	`\$`	$
`\textdaggerdbl`	‡	`\^{}`	^
`\textless`	<	`\_`	_
`\textgreater`	>		

1.7.1 The `geometry` Package

Some students are unfortunately saddled with misguided supervisors or committee members whom believe that documents should have the narrowest margins possible. The `geometry` package is helpful in such cases. Using this package is as simple as:

```
\usepackage[margin=1.25in,top=1.25in,bottom=1.25in]{geometry}
```

You should check the package's documentation for more complex uses.

1.7.2 Changing Page Layout Values By Hand

There are some miserable students with requirements for page layouts that vary throughout the document. Unfortunately the `geometry` can only be specified once, in the document's preamble. Such miserable students must set LaTeX's layout parameters by hand:

```
\setlength{\topmargin}{-.75in}
\setlength{\headsep}{0.25in}
\setlength{\headheight}{15pt}
\setlength{\textheight}{9in}
\setlength{\footskip}{0.25in}
\setlength{\footheight}{15pt}


% The *sidemargin values are relative to 1in; so the following
% results in a 0.75 inch margin
\setlength{\oddsidemargin}{-0.25in}
\setlength{\evensidemargin}{-0.25in}
\setlength{\textwidth}{7in}     % 1.1in margins (8.5-2*0.75)
```

These settings necessarily require assuming a particular page height and width; in the above, the setting for `\textwidth` assumes a US Letter with an 8.5" width. The `geometry` package simply uses the page height and other specified values to derive the other layout values. The `layout` package provides a handy `\layout` command to show the current page layout parameters.

7

1.7.3 Making Temporary Changes to Page Layout

There are occasions where it becomes necessary to make temporary changes to the page width, such as to accomodate a larger formula. The `chngpage` package provides an `adjustwidth` environment that does just this. For example:

```
% Expand left and right margins by 0.75in
\begin{adjustwidth}{−0.75in}{−0.75in}
% Must adjust the perceived column width for LaTeX to get with it.
\addtolength{\columnwidth}{1.5in}
\[ an extra long math formula \]
\end{adjustwidth}
```

1.8 Keeping Track of Versions with Revision Control

Software engineers have used Revision control system (RCS) to track changes to their software systems for decades. These systems record the changes to the source code along with context as to why the change was required. These systems also support examining and reverting to particular revisions from their system's past.

An RCS can be used to keep track of changes to things other than source code, such as your dissertation. For example, it can be useful to know exactly which revision of your dissertation was sent to a particular committee member. Or to recover an accidentally deleted file, or a badly modified image. With a revision control system, you can tag or annotate the revision of your dissertation that was sent to your committee, or when you incorporated changes from your supervisor.

Unfortunately current revision control packages are not yet targetted to non-developers. But the Subversion project's TortoiseSVN[8] has greatly simplified using the Subversion revision control system for Windows users. You should consult your local geek.

A simpler alternative strategy is to create a GoogleMail account and periodically mail yourself zipped copies of your dissertation.

1.9 Recommended Packages

The real strength to LaTeX is found in the myriad of free add-on packages available for handling special formatting requirements. In this section we list some helpful packages.

1.9.1 Typesetting

enumitem: Supports pausing and resuming enumerate environments.

ulem: Provides two new commands for striking out and crossing out text (`\sout{text}` and `\xout{text}` respectively) The package should likely be used as follows:

---

[8]http://tortoisesvn.net/docs/release/TortoiseSVN_en/

```
\usepackage[normalem,normalbf]{ulem}
```

to prevent the package from redefining the emphasis and bold fonts.

chngpage: Support changing the page widths on demand.

mhchem: Support for typesetting chemical formulae and reaction equations.

Although not a package, the `latexdiff`[9] command is very useful for creating changebar'd versions of your dissertation.

1.9.2 Figures, Tables, and Document Extracts

pdfpages: Insert pages from other PDF files. Allows referencing the extracted pages in the list of figures, adding labels to reference the page from elsewhere, and add borders to the pages.

subfig: Provides for including subfigures within a figure, and includes being able to separately reference the subfigures. This is a replacement for the older `subfigure` environment.

rotating: Provides two environments, sidewaystable and sidewaysfigure, for typesetting tables and figures in landscape mode.

longtable: Support for long tables that span multiple pages.

tabularx: Provides an enhanced tabular environment with auto-sizing columns.

ragged2e: Provides several new commands for setting ragged text (e.g., forms of centered or flushed text) that can be used in tabular environments and that support hyphenation.

1.9.3 Bibliography Related Packages

bibunits: Support having per-chapter bibliographies.

footbib: Cause cited works to be rendered using footnotes.

bibentry: Support placing the details of a cited work in-line.

custom-bib: Generate a custom style for your bibliography.

---

[9]http://www.ctan.org/tex-archive/support/latexdiff/

1.10 Moving On

At this point, you should be ready to go. Other handy web resources:

- Common TeX Archive Network (CTAN)[10] is *the* comprehensive archive site for all things related to TeX and LaTeX. Should you have some particular requirement, somebody else is almost certainly to have had the same requirement before you, and the solution will be found on CTAN. The links to various packages in this document are all to CTAN.

- An online reference to LaTeX commands[11] provides a handy quick-reference to the standard LaTeX commands.

- The list of Frequently Asked Questions about TeX and LaTeX[12] can save you a huge amount of time in finding solutions to common problems.

- The teTeX documentation guide[13] features a very handy list of the most useful packages for LaTeX as found in CTAN.

- The `color`[14] package, part of the graphics bundle, provides handy commands for changing text and background colours. Simply changing text to various levels of grey can have a very dramatic effect.

- If you're really keen, you might want to join the TeX Users Group[15].

---

[10] http://www.ctan.org

[11] http://www.ctan.org/get/info/latex2e-help-texinfo/latex2e.html

[12] http://www.tex.ac.uk/cgi-bin/texfaq2html?label=interruptlist

[13] http://www.tug.org/tetex/tetex-texmfdist/doc/

[14] http://www.ctan.org/tex-archive/macros/latex/required/graphics/grfguide.pdf

[15] http://www.tug.org

Bibliography

[1] R. Bringhurst. *The Elements of Typographic Style*. Hartley & Marks, 2.5 edition, 2002. ISBN 0881791326. → pages 6

[2] G. Kiczales, J. Lamping, A. Mendhekar, C. Maeda, C. Lopes, J.-M. Loingtier, and J. Irwin. Aspect-oriented programming. In *Proceedings of the European Conference on Object-Oriented Programming (ECOOP)*, volume 2591 of *Lecture Notes in Computer Science*, pages 220–242, 1997. → pages 3, 4

[3] L. Lamport. *LaTeX: A Document Preparation System*. Addison-Wesley, 2 edition, 1994. ISBN 0201529831. → pages 1, 4

Appendix A

Supporting Materials

This would be any supporting material not central to the dissertation. For example:

- additional details of methodology and/or data;

- diagrams of specialized equipment developed.;

- copies of questionnaires and survey instruments.