

**TARGETED NEXT-GENERATION SEQUENCING 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

© Shyong Quin Yap, 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

Abstract	ii
Preface	iii
Table of Contents	iv
List of Tables	vi
List of Figures	vii
Glossary	viii
Acknowledgments	ix
1 Introduction	1
1.1 Cancer as an (Epi)genetic Disease	1
1.2 The Era of Precision Oncology	1
1.2.1 Next-Generation Sequencing Technologies in the Clinic	2
1.2.2 Clinical Application of Bioinformatics Pipelines	2
1.2.3 Molecular Targeted Therapies	2
1.2.4 Cancer Pharmacogenomics	2
1.3 Challenges in Clinical Applications of Next-Generation Sequencing	2
1.3.1 Formalin-Fixed Clinical Specimens	2
1.3.2 Tumour-Only Profiling	2
1.4 BibTeX	2
1.4.1 Describing References	2
1.4.2 Citing References	3
1.4.3 Formatting Cited References	3
1.5 Typesetting Tables	4
1.6 Figures, Graphics, and Special Characters	5

1.7	Special Characters and Symbols	6
1.8	Changing Page Widths and Heights	6
1.8.1	The geometry Package	6
1.8.2	Changing Page Layout Values By Hand	7
1.8.3	Making Temporary Changes to Page Layout	7
1.9	Keeping Track of Versions with Revision Control	7
1.10	Recommended Packages	8
1.10.1	Typesetting	8
1.10.2	Figures, Tables, and Document Extracts	8
1.10.3	Bibliography Related Packages	9
1.11	Moving On	9
	Bibliography	11
	A Supporting Materials	12

List of Tables

Table 1.1	Available <code>cite</code> variants; the exact citation style depends on whether the bibliography style is numeric or author-year.	4
Table 1.2	Useful \LaTeX symbols	6

List of Figures

Figure 1.1	Proof of L ^A T _E X's amazing abilities	5
------------	--	---

Glossary

This glossary uses the handy `acroynm` 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 gratitude 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

1.1 Cancer as an (Epi)genetic Disease

Cancer is a group of disorders defined by abnormal cell growth. Although fundamentally known to arise from genetic mutations, the disease paradigm has expanded to include aberrant epigenetic mechanisms as a contributing factor to oncogenesis. The understanding of cancer pathogenesis has expanded been increasing over the years and a disorder that was fundamentally known to arise from genetic mutations this group of disorders which have been fundamentally known Cancer has been fundamentally known as a genetic disease defined by abnormal proliferation of cells. Our understanding of cancer pathogenesis has been expanding Although the understanding of cancer pathogenesis has been expanding, Cancer has been fundamentally known as a genetic disease.

1.2 The Era of Precision Oncology

The era of precision oncology has been revolutionized by NGS technologies and bioinformatics algorithms etc. This results in the use of molecular targeted therapies to treat patients and PGx info to administer chemo.

1.2.1 Next-Generation Sequencing Technologies in the Clinic

1.2.2 Clinical Application of Bioinformatics Pipelines

1.2.3 Molecular Targeted Therapies

1.2.4 Cancer Pharmacogenomics

1.3 Challenges in Clinical Applications of Next-Generation Sequencing

1.3.1 Formalin-Fixed Clinical Specimens

1.3.2 Tumour-Only Profiling

1.4 BibTeX

One of the primary benefits of using L^AT_EX is its companion program, BibT_EX, for managing bibliographies and citations. Managing bibliographies has three parts: (i) describing references, (ii) citing references, and (iii) formatting cited references.

1.4.1 Describing References

BibT_EX defines a standard format for recording details about a reference. These references are recorded in a file with a `.bib` extension. BibT_EX 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*.¹ 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 simplifies remembering how to make references. For example:

`last-name`–`year`–`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`.

¹Note that the citation keys are different from the reference identifiers as described in ??.

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

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

1.5 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 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
\end{tabular}
\end{table}

```

L^AT_EX Rocks!

Figure 1.1: Proof of L^AT_EX’s amazing abilities

```
% ... 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.

1.6 Figures, Graphics, and Special Characters

Most L^AT_EX beginners find figures to be one of the more challenging topics. In L^AT_EX, 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 T_EX 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}
```

Graduate and Postdoctoral Studies (GPS) generally prefers shortened single-line captions in the LoF: multiple-line captions are a bit unwieldy.

Table 1.2: Useful L^AT_EX symbols

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

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

See Keith Reckdahls excellent guide for more details, *Using imported graphics in LaTeX2e*⁴.

1.7 Special Characters and Symbols

\LaTeX appropriates many common symbols for its own purposes, with some used for commands (i.e., $\backslash \{ \}$ & $\%$) and mathematics (i.e., $\_), and others are automatically transformed into typographically-preferred forms (i.e., $-''$) or to completely different forms (i.e., $\langle \rangle$). 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/>

1.8 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!

1.8.1 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.

⁴<http://www.ctan.org/tex-archive/info/epslatex.pdf>

1.8.2 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.

1.8.3 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 `chngepage` 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}
```

1.9 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 TortoiseSVN⁵ 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.

1.10 Recommended Packages

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

1.10.1 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:

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

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

chnngpage: Support changing the page widths on demand.

mhchem: Support for typesetting chemical formulae and reaction equations.

Although not a package, the `latexdiff`⁶ command is very useful for creating changebar'd versions of your dissertation.

1.10.2 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.

⁵<http://tortoisetsvn.net/docs/release/TortoiseSVN.en/>

⁶<http://www.ctan.org/tex-archive/support/latexdiff/>

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.

1.10.3 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.

1.11 Moving On

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

- Common T_EX Archive Network (CTAN)⁷ is *the* comprehensive archive site for all things related to T_EX and L^AT_EX. 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 L^AT_EX commands⁸ provides a handy quick-reference to the standard L^AT_EX commands.
- The list of Frequently Asked Questions about T_EX and L^AT_EX⁹ can save you a huge amount of time in finding solutions to common problems.
- The teT_EX documentation guide¹⁰ features a very handy list of the most useful packages for L^AT_EX as found in CTAN.

⁷<http://www.ctan.org>

⁸<http://www.ctan.org/get/info/latex2e-help-texinfo/latex2e.html>

⁹<http://www.tex.ac.uk/cgi-bin/texfaq2html?label=interruptlist>

¹⁰<http://www.tug.org/tetex/tetex-texmfdist/doc/>

- The `color`¹¹ 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 T_EX Users Group¹².

¹¹<http://www.ctan.org/tex-archive/macros/latex/required/graphics/grfguide.pdf>

¹²<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 2, 3, 4
- [3] L. Lamport. *LaTeX: A Document Preparation System*. Addison-Wesley, 2 edition, 1994. ISBN 0201529831. → pages 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.