

# The PEP.db User Guide

Renan Sauteraud\*

November 14, 2013

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Peptide collections</b>	<b>1</b>
2.1	Information . . . . .	1
2.2	Structure . . . . .	2
2.3	pep_hxb2 . . . . .	2
2.4	pep_hxb2JPT . . . . .	2
2.5	pep_mac239 . . . . .	2
2.6	pep_m239smE543 . . . . .	2
<b>3</b>	<b>Loading the collections</b>	<b>2</b>
<b>4</b>	<b>Create collections</b>	<b>3</b>
4.1	Requirements . . . . .	3
4.2	Constructors . . . . .	3

## 1 Introduction

PEP.db is an R package that stores peptide collections and provide functions for their analysis. It is intended to be used in conjunction with other packages for peptide analysis and visualisation: HIV.db, pepStat and Pviz.

As with any R package, it should first be loaded in the session

```
library(PEP.db)
```

## 2 Peptide collections

### 2.1 Information

For each peptide, the following collections display information about the position relative to the reference sequence, the alignment, the trimmed alignment, the zSums for the physicochemical properties as well as

---

\*rsautera@fhcrc.org

the clades where they can be found.

## 2.2 Structure

The datasets in this package are **RangedData** objects. For more information about the class, its accessors and setters, please refer to **IRanges** documentation.

## 2.3 pep\_hxb2

This collection is based on the alignment of the reference HIV sequence hxb2 and seven subtypes (clades) A, B, C, D, M, CRF01 and CRF02.

## 2.4 pep\_hxb2JPT

This collection adds a few more clades to the previous one: CM244, CON\_01\_AE, LAI\_A04321 and MN\_DD328842.

## 2.5 pep\_mac239

This collection is for SIV, with the clades mac239 and E660.

## 2.6 pep\_m239smE543

This collection is for SIV, with the clades mac239 and E543.

## 3 Loading the collections

The peptide collections can be loaded like any other R data.

```
data(pep_hxb2)
head(pep_hxb2)

## RangedData with 6 rows and 10 value columns across 1 space
##           space  ranges |           names           aligned
##           <factor> <IRanges> | <character>           <factor>
## MRVKETQMNWPNLWK   gp160  [1, 16] | MRVKETQMNWPNLWK MRVKETQMNWPNL----WK
## MRVMGIQKNYPPLLWR   gp160  [1, 16] | MRVMGIQKNYPPLLWR MRVMGIQKNYPPLL----WR
## MRVMGIQRNCQHLWR    gp160  [1, 16] | MRVMGIQRNCQHLWR MRVMGIQRNCQHL----WR
## MRVKGIRKKNYQHLWR   gp160  [1, 16] | MRVKGIRKKNYQHLWR MRVKGIRKKNYQHL----WR
## MRVRGILRNWQQWWI    gp160  [1, 16] | MRVRGILRNWQQWWI MRVRGILRNWQQW----WI
## MRVRGIERNYQHLWR    gp160  [1, 16] | MRVRGIERNYQHLWR MRVRGIERNYQHL----WR
##           trimmed  seqNb    clade      z1      z2
##           <factor> <integer> <character> <numeric> <numeric>
## MRVKETQMNWPNLWK MRVKenWPNL----WK          1      CRF01      -3.14      9.87
## MRVMGIQKNYPPLLWR MRVMgnYPLL----WR          1      CRF02     -13.12      3.19
## MRVMGIQRNCQHLWR MRVMgnCQHL----WR          1          A       1.65      4.17
## MRVKGIRKKNYQHLWR MRVKgnYQHL----WR          1          B       3.95      9.78
```

##	MRVRGILRNWQQWWI	MRVRgnWQQW----	WI	1	C	-11.42	10.26
##	MRVRGIERNYQHLWR	MRVRgnYQHL----	WR	1	D	6.00	10.76
##		z3	z4	z5			
##		<numeric>	<numeric>	<numeric>			
##	MRVKETQMNPWLWK	-8.14	6.72	1.36			
##	MRVMGIQKNYPLLWR	-11.96	7.09	0.76			
##	MRVMGIQRNCQHLWR	-11.22	10.31	-2.99			
##	MRVKGIRKNYQHLWR	-18.51	12.55	-0.87			
##	MRVRGILRNWQQWWI	-16.19	10.33	-2.37			
##	MRVRGIERNYQHLWR	-17.14	8.52	-1.91			

## 4 Create collections

While the package comes with datasets for HIV and SIV. It is possible to create new collections for different organisms or proteins.

### 4.1 Requirements

Minimum information required to create a peptide collection:

- start
- end or width
- peptide sequence

### 4.2 Constructors

PEP.db includes a constructor that will generate the appropriate object from a `data.frame` that has at least a column for each requirement and any additional information.

The objects being `RangedData`, it is naturally possible to use an `IRanges` object to specify the ranges of the peptides.