

# The PEP.db User Guide

Renan Sauteraud\*

January 24, 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	pep_hxb2 . . . . .	1
2.3	pep_hxb2JPT . . . . .	1
2.4	pep_mac239 . . . . .	1
<b>3</b>	<b>Loading the collections</b>	<b>2</b>

## 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 collection display information about the position relative to the reference sequence, the alignment, the trimmed alignment, the zSums for the physicochemical properties as well as the clades where they can be found.

### 2.2 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.

---

\*rsautera@fhcrc.org

## 2.3 pep\_hxb2JPT

This collection adds a few more clades to the previous one: CM244, CON\_01\_AE, LALA04321 and MN\_DD328842.

## 2.4 pep\_mac239

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

## 3 Loading the collections

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

```
> data(pep_hxb2)
> pep_hxb2
```

RangedData with 1423 rows and 10 value columns across 1 space

	space	ranges		names	aligned	
	<factor>	<IRanges>		<character>	<factor>	
MRVKETQMNWPNLWK	gp160	[1, 16]		MRVKETQMNWPNLWK	MRVKETQMNWPNL----WK	
MRVMGIQKNYPLLWR	gp160	[1, 16]		MRVMGIQKNYPLLWR	MRVMGIQKNYPLL----WR	
MRVMGIQRNCQHLWR	gp160	[1, 16]		MRVMGIQRNCQHLWR	MRVMGIQRNCQHL----WR	
MRVKGIRKNYQHLWR	gp160	[1, 16]		MRVKGIRKNYQHLWR	MRVKGIRKNYQHL----WR	
MRVRGILRNWQQWWI	gp160	[1, 16]		MRVRGILRNWQQWWI	MRVRGILRNWQQW----WI	
MRVRGIERNYQHLWR	gp160	[1, 16]		MRVRGIERNYQHLWR	MRVRGIERNYQHL----WR	
MRVMGIQRNWQHLWR	gp160	[1, 16]		MRVMGIQRNWQHLWR	MRVMGIQRNWQHL----WR	
KETQMNWPNLWKWGT	gp160	[4, 19]		KETQMNWPNLWKWGT	KETQMNWPNL----WKWGT	
MGIQKNYPLLWRWGT	gp160	[4, 19]		MGIQKNYPLLWRWGT	MGIQKNYPLL----WRWGT	
...	...	...	...	...	...	
VRAILNIPRRIRQGL	gp160	[837, 851]		VRAILNIPRRIRQGL	VRAILNIPRRIRQGL	
GRAILNIPRRIRQGL	gp160	[837, 851]		GRAILNIPRRIRQGL	GRAILNIPRRIRQGL	
ILHIPRRIRQGFERA	gp160	[840, 854]		ILHIPRRIRQGFERA	ILHIPRRIRQGFERA	
ILHIPRRIRQGLERA	gp160	[840, 854]		ILHIPRRIRQGLERA	ILHIPRRIRQGLERA	
IRNIPRRIRQGFEEA	gp160	[840, 854]		IRNIPRRIRQGFEEA	IRNIPRRIRQGFEEA	
ILNIPRRIRQGLERA	gp160	[840, 854]		ILNIPRRIRQGLERA	ILNIPRRIRQGLERA	
IPRRIRQGFERALL	gp160	[843, 856]		IPRRIRQGFERALL	IPRRIRQGFERALL	
IPRRIRQGFEEAALL	gp160	[843, 856]		IPRRIRQGFEEAALL	IPRRIRQGFEEAALL	
IPRRIRQGLERALL	gp160	[843, 856]		IPRRIRQGLERALL	IPRRIRQGLERALL	
	trimmed	seqNb		z1sum	z2sum	z3sum
	<factor>	<integer>		<numeric>	<numeric>	<numeric>
MRVKETQMNWPNLWK	MRVKenWPNL----WK	1		-4.16	-0.27	-2.17
MRVMGIQKNYPLLWR	MRVMgnYPLL----WR	1		-0.67	-10.74	-9.44
MRVMGIQRNCQHLWR	MRVMgnCQHL----WR	1		0.07	-14.24	-17.48
MRVKGIRKNYQHLWR	MRVKgnYQHL----WR	1		1.40	-7.86	-10.03
MRVRGILRNWQQWWI	MRVRgnWQQW----WI	1		-3.65	-3.35	0.55
MRVRGIERNYQHLWR	MRVRgnYQHL----WR	1		7.65	-10.82	2.52
MRVMGIQRNWQHLWR	MRVMgnWQHL----WR	1		-4.03	11.41	-0.83

KETQMNWPNLWKWGT	KenWPNL----	WKWGT	2	-10.66	-0.97	3.69
MGIQKNYPLLWRWGT	MgnYPLL----	WRWGT	2	-2.98	-12.68	-9.47
...	...	...	...	...	...	...
VRAILNIPRRIRQGL	VRAILNIPRRIRQGL		276	-36.72	-1.82	-7.29
GRAILNIPRRIRQGL	GRAILNIPRRIRQGL		276	15.29	0.38	-3.85
ILHIPRRIRQGFERA	ILHIPRRIRQGFERA		277	12.14	6.13	-1.39
ILHIPRRIRQGLERA	ILHIPRRIRQGLERA		277	3.42	7.47	-0.22
IRNIPRRIRQGFEEA	IRNIPRRIRQGFEEA		277	14.90	-1.16	4.64
ILNIPRRIRQGLERA	ILNIPRRIRQGLERA		277	-2.97	-1.92	-5.38
IPRRIRQGFERALL	IPRRIRQGFERALL		278	27.97	-2.55	5.05
IPRRIRQGFEEALL	IPRRIRQGFEEALL		278	-6.69	8.72	-1.78
IPRRIRQGLERALL	IPRRIRQGLERALL		278	3.39	-15.01	-2.39
	z4sum	z5sum		clade		
	<numeric>	<numeric>		<character>		
MRVKETQMNWPNLWK	-4.02	2.53		CRF01		
MRVMGIQKNYPLLWR	-10.92	6.27		CRF02		
MRVMGIQRNCQHLWR	-12.50	6.22		A		
MRVKGIRKNYQHLWR	-4.99	8.89		B		
MRVRGILRNWQQWWI	-5.78	7.85		C		
MRVRGIERNYQHLWR	-13.20	0.03		D		
MRVMGIQRNWQHLWR	1.21	-2.59		M		
KETQMNWPNLWKWGT	-4.67	2.30		CRF01		
MGIQKNYPLLWRWGT	0.30	0.04		CRF02		
...	...	...		...		
VRAILNIPRRIRQGL	7.77	-7.83		D		
GRAILNIPRRIRQGL	-2.41	2.20		M,CRF02		
ILHIPRRIRQGFERA	-2.43	-1.45		A		
ILHIPRRIRQGLERA	-7.40	2.25		B,CRF01		
IRNIPRRIRQGFEEA	-4.71	2.38		C		
ILNIPRRIRQGLERA	-9.22	5.74		M,D,CRF02		
IPRRIRQGFERALL	-13.52	7.94		A		
IPRRIRQGFEEALL	-1.95	-0.29		C		
IPRRIRQGLERALL	-10.26	0.48		M,B,D,CRF01,CRF02		