Convert the species of a spectral reference library with dialects

Madeleine J Otway

2018-02-13

Package

dialects 0.0.0.56

Contents

Convert the species of a spectral reference library with dialects

```
library(dialects)
library(knitr)
library(kableExtra)
```

1 Introduction to dialects

This package has been designed to help convert spectral reference libraries (SRLs) from one species to another. SRLs are converted by matching peptides from an *in silico* trypsin digestion of a protein sequence database (fasta file). Peptides are only matched if they have full sequence identity (i.e. the peptides are identical). This package is currently compatible with PeakView, OneOmics and OpenSWATH formatted SRLs. The latest version of the package can be found at: https:github.com/madeleineotway/dialects. Any enquiries should be sent to motway@cmri.org.au

2 Import a protein sequence database of desired species

To convert between species, you need to start with a protein sequence database of the desired species for your SRL. This function will allow you to import said database for *in silico* digestion in the next step.

To download a protein sequence database, head to www.uniprot.org/. The database must be in fasta format (specifically .fasta only) or the function will not work. This function will work with SwissProt and TrEMBL formatted protein sequences. Once downloaded, run the function to import the fasta file.

```
human_fasta <- system.file("extdata",
                           "human_proteome_example.fasta",
                           package = "dialects")
human_proteome <- import.fasta(human_fasta)
```

accession

name

sequence

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

MFHLRTCAAKLRPLTASQTVKTFSQNRPAAARTFQ QIRCYSAPVAAEPFLSGTSS-NYVEEMYCAWLENPK SVHKSWDIFFRNTNAGAPPGTAYQSPLPLSRGSLA AVAHAQS-LVEAQPNVDKLVEDHLAVQSLIRAYQIR GHHVAQLDPLGILDADLDSSVPADIISSTDKLGFY GLDESDLDKVFHLPTTTFIGGQESALPLREIIRRL EMAYCQHIGVEFMFINDLEQCQWIRQK-FETPGIMQ FTNEEKRTLLARLVRSTRFEEFLQRKWSSEKRFGL EGCEVLIPALKTI-IDKSSENGVDYVIMGMPHRGRL NVLANVIRKELEQIFCQFDSKLEAADEGSGDVKYH

LGMYHRRINRVTDRNITLSLVANPSHLEAADPVVM GKTKAEQFYCGDTEGKKVMSILL-
HGDAAFAGQGIV YETFHLSDLPSYTTHGTVHVVVNNQIGFTTDPRMA RSSPYPT-
DVARVVNAPIFHVNSDDPEAVMYVCKVA AEWRSTFHKDVVVDLVCYRRNGH-
NEMDEPMFTQPL MYKQIRKQKPVLQKYAELLVSQGVVNQPEYEEEIS KYDKICEEA-
FARSKDEKILHIKHWLDSPWPGFFTL DGQPRSMSCPSTGLTEDILTHIGNVASSVPVENFT
IHGGLSRILKTRGEMVKNRTVDWALAEYMAFGSLL KEGIHIRLSGQDVERGTFSHRHHVL-
HDQNVDKRTC IPMNHLWPNQAPYTVCNSSLSEYGVLGFELGFAMA SPNALVL-
WEAQFGDFHNTAQCIIDQFICPGQAKWV RQNGIVLLLPHGMEGMGPEHSSARPER-
FLQMCNDD PDVLPDLKEANFDINQLYDCNWVVVNCSTPGNFFH VLRRQILLPFRK-
PLIIFTPKSLLRHPEARSSFDEM LPGTHFQRVIPEDGPAAQNPENVKRLLFCTGKVYY
DLTRERKARDMVGQVAITRIEQLSPFPFDLLLKEV QKYPNAELAWCQEEHKNQGYY-
DYVKPRLRTTISRA KPVWYAGRDPAAAPATGNKKTHLTELQRLLDTAFD LDVFKNFS

sp|P50993|AT1A2_HUMAN

Sodium/potassium-transporting ATPase subunit alpha-2 OS=Homo sapiens GN=ATP1A2
PE=1 SV=1

MGRGAGREYSPAATTAENGGGKKKQKEKELDELKK EVAMDDHKLSLDELGRKYQVDL-
SKGLTNQRAQDVL ARDGPNALTPPPTTPEWVKFCRQLFGGFSILLWIG AILCFLAY-
GIQAAMEDEPSNDNLYLGVVLAAVVIV TGCFSYYQEAKSSKIMDSFKNMVPQQALVIREGEK
MQINAEEVVVGDLVEVKGGDRVPADLRIISSHGCK VDNSSLTGESEPQTRSPEFTHEN-
PLETRNICFFST NCVEGTARGIVIATGDRTVMGRIATLASGLEVGRT PIAMEIEHFIQL-
ITGVAVFLGVSFFVLSLILGYSW LEAVIFLIGIIVANVPEGLLATVTVCLTLTAKRMA RKN-
CLVKNLEAVETLGSTSTICSDKTGTLTQNRMT VAHMWFDNQIHEADTTEDQSGATFD-
KRSPTWTALS RIAGLCNRAVFKAGQENISVSKRDTAGDASESALL KCIELSCGSVRKM-
RDRNPKVAEIPFNSTNKYQLSI HEREDSPQSHVLVMKGAPERILDRCSTILVQGKEI
PLDKEMQDAFQNAYMELGGLGERVLGFCQLNLPSG KFPRGFKFDTDELNFPTEKL-
CFVGLMSMIDPPRAA VPDAVGKCRSAGIKVIMVTGDHPITAKAIAKGVGI ISEGNETVEDI-
AARLNIPMSQVNPREAKACVVHGS DLKDMTSEQLDEILKNHTEIVFARTSPQQKLIIVE
GCQRQGAIVAVTGDGVNDSPALKKADIGIAMGISG SDVSKQAADMILLDDNFASIVTGVEE-
GRLIFDNLK KSIAYTLTSNIPEITPFLLFIIANIPLPLGTVTIL CIDLGTDMVPAISLAYEAAES-
DIMKRQPRNSQTDK LVNERLISMAYGQIGMIQALGGFFTYFVILAENGF LPSRLLGIRLDWD-
DRTMNDLEDSYGQEWTYEQRKV VEFTCHTAFFASIVVVQWADLIICKTRRNSVFQQG
MKNKILIFGLLEETALAAFLSYCPGMGVALRMYPL KVTWWFCAFPYSLLIFIYDEVRKLIL-
RRYPGGWVE KETYY

sp|Q9Y4I1|MYO5A_HUMAN

Unconventional myosin-Va OS=Homo sapiens GN=MYO5A PE=1 SV=2

MAASELYTKFARVWIPDPEEVWKSAELLKDYKPGD KVLLLHLEEGKDLEYHLDP-
KTKELPHLRNPDILVG ENDLTALSYLHEPAVLHNLRVRFIDSKLIYTYCGI VLVAIN-
PYEQLPIYGEDIINAYSGQNMGDMDPHIF AVAEEAYKQMARDERNQSIIVSGESGAGK-
TVSAKY AMRYFATVSGSASEANVEEKVLASNPIMESIGNAK TTRNDNSSRFGKYIEIGFD-
KRYRIIGANMRTYLLE KSRVVFQAEEERNYHIFYQLCASAKLPEFKMLRLG NAD-
NFNYTKQGGSPVIEGVDDAKEMAHTRQACTLL GISESHQMGIFRILAGILHLGNVGFT-
SRDADSCTI PPKHEPLCIFCELMGVDYEEMCHWLCHRKLATATE TYIKPISKLQAT-
NARDALAKHIYAKLFNWIVDNVN QALHSAVKQHSFIGVLDIYGFETFEINSFEQFCIN
YANEKLQQQFNMHVFKLEQEEYMKEQIPWTLIDFY DNQPCINLIESKLGILDLLDEECKMP-
KGTDDTWAQ KLYNTHLNKCALFEKPRLSNKAFIIQHFADKVEYQ CEGFLEKNKDTVFEE-
QIKVLKSSKFKMLPELFQDD EKAISPTSATSSGRTPLTRTPAKPTKGRPGQMAKE
HKKTVGHQFRNSLHLLMETLNATTPHYVRCIKPND FKFPFTFDEKRAVQQLRACGVLE-
TIRISAAGFPSR WTYQEFFSRYRVLMKQKDVLSDRKQTCKNVLEKLI LDKDKYQFGK-
TKIFFRAGQVAYLEKLRADKLRAAC IRIQKTIRGWLLRKKYLRMRKAAITMQRYVR-

GYQA RCYAKFLRRTKAATIIQKYWRMYVVRRRYKIRRAA TIVLQSYLRGFLARN-
RYRKILREHKAVIIQKRVRG WLARTHYKRSMHAIIYLQCCFRRMMAKRELKKLKI
EARSVERYKKLHIGMENKIMQLQRKVDEQNKDYKC LVEKLTNLEGIYNSETEKLRSDLER-
LQLSEEEAKV ATGRVLSLQEEIAKLRKDLEQTRSEKKCIEEHADR YKQETEQLVSNL-
KEENTLLKQEKEALNHRIVQQAK EMTETMEKKLVEETKQLELDLNDERLRYQNLL-
NEF SRLEERYDDLKEEMTLMVHVPKPGHKRTDSTHSSN ESEYIFSSEIAEMEDIP-
SRTEEPSEKKVPLDMSLF LKLQKRVTELEQEKQVMQDELDRKEEQVLRSKAKE
EERPQIRGAELEYESLKRQELESENKKLKNELNEL RKALSEKSAPEVTAPGAPAYRVLME-
QLTSVSEELD VRKEEVLILRSQLVSQKEAIQPKDDKNTMTDSTIL LEDVQKMKD-
KGEIAQAYIGLKETNRSSALDYHELN EDGELWLVYEGLKQANRLLESQLQSQKRSHENEAE
ALRGEIQSLKEENNRQQQLLAQNLQLPPEARIEAS LQHEITRLTNENLDLMEQLEKQDK-
TVRKLKKQLKV FAKKIGELEVGQMENISPGQIIDEPIRPVNIPRKE KDFQGMLEYKKED-
EQKLVKNLILELKPRGVAVNLI PGLPAYILFMCVRHADYLNDDQKVRSLLTSTINSI
KKVLKKRGDDFETVSFWLSNTCRFLHCLKQYSGEE GFMKHNTSRQNEHCLTNFD-
LAEYRQVLSDLAIQIY QQLVRVLENILQPMIVSGMLEHETIQGVSGVKPTG LRKRTSSI-
ADEGTYTLDSILRQLNSFHSVMCQHGM DPELIKQVVKQMFYIIGAITLNNLLLRKDMC-
SWSK GMQIRYNVSQLEEWLRDKNLMNSGAKETLEPLIQA AQLLQVKKKTDDDAEAICSM-
CNALTTAQIVKVLNL YTPVNEFEERVSVSFIRTIQMRLRDRKDSPQLLMD AKHIFPVTF-
PFNPSSLALETIQIPASLGLGFISRV

sp|P01023|A2MG_HUMAN

Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3

MGKNKLLHPSLVLLLLVLLPTDASVSGKPQYMVLV PSLLHTETTEKGCVLLSYL-
NETVTVSASLESVRGN RSLFTDLEAENDVLHCVAFAVPKSSSNEEVMFLTV QVKG-
PTQEFKKRTTVMVKNEDSLVFVQTDKSIYKP GQTVKFRVVSMDENFHPLNELI-
PLVYIQDPKGNRI AQWQSFQLEGGLKQFSFPLSSEPFQGSYKVVVQKK SGGRTEH-
PFTVEEFVLPKFEVQVTVPKIITILEEE MNVSVCGLYTYGKPVPGHVTVSICRKYSDASD-
CHG EDSQAFCEKFSGQLNSHGCFYQQVKTKVFQLKRKE YEMKLHTEAQIQEEGTVVELT-
GRQSSEITRTITKL SFVKVDSHFRQGIPFFGQVRLVDGKGVPIPNKVIF IRGNEANYYS-
NATTDEHGLVQFSINTTNVMGTSLT VRVNYKDRSPCYGYQWVSEEHEEAHHTAYLVF-
SPS KSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLGL KKLSFYYLIMAKGGIVRT-
GTHGLLVKQEDMKGHFS ISIPVKSDIAPVARLLIYAVLPTGDVIGDSAKYDV EN-
CLANKVDLSFSPSQSLPASHAHLRVTAAPQSVC ALRAVDQSVLLMKPDAELSASSVYN-
LLPEKDLTGF PGPLNDQDNEDCINRHNVYINGITYTPVSSTNEKD MYSFLEDMGLKA-
FTNSKIRKPKMCPQLQQYEMHGP EGLRVGFYESDVMGRGHARLVHVEEPHTETVRKYF
PETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAF CLSEDAGLGISSTASLRAFQPF-
FVELTMPYSVIRG EAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVE KEQAPHCICAN-
GRQTVSWAVTPKSLGNVNFTVSAE ALESQELCGTEVPSVPEHGRKDTVIKPLLVEPEGL
EKETTFNSLLCPSGGEVSEELSLKLPPNVVEESAR ASVSVLGDILGSAMQNTQNL-
LQMPYGCGEQNMVLF APNIYVLDYLNETQQLTPEIKSKAIGYLNTGYQRQ LNYKHY-
DGSYSTFGERYGRNQGNTWLTAFVLKTFA QARAYIFIDEAHITQALIWLSQRQKD-
NGCFRSSGS LLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPV VRNALFCLESAWK-
TAQEGDHGSHVYTKALLAYAFA LAGNQDKRKEVLKSLNEEAVKKDNSVHWERPQKPK
APVGHFYEPQAPSAEVEMTSYVLLAYLTAQPAPTS EDLTSATNIVKWITKQQNAQG-
GFSSTQDTVVALHA LSKYGAATFTRTGKAAQVTIQSSGTFSSKFQVDNN NRLL-
LQQVSLPELPGEYSMKVTGEGCVYLQTSLKY NILPEKEEFPFALGVQTLPQTCDEPKAHTS-
FQISL SVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLE RSNHVSRTEVSSNHVLIYLD-
KVSNQTLSLFFTVLQ DVPVRDLKPAIVKVYDYYETDEFAIAEYNAPCSKD LGNA

sp|P17661|DESM_HUMAN

Desmin OS=Homo sapiens GN=DES PE=1 SV=3

MSQAYSSSQRVSSYRRTFGGAPGFPLGSPLSSPVF PRAGFGSKGSSSSVTSRVYQVS-
RTSGGAGGLGSLR ASRLGTTRTPSSYGAGELLDFSLADAVNQEFLTTR TNEKVELQEL-
NDRFANYIEKVRFLEQQNAALAAEV NRLKGREPTRVAELYEEELRELRRQVEVLT-
NQRAR VDVERDNLLDDLQRLKAKLQEEIQLKEEAENNLAA FRADVDAATLARI-
DLERRIESLNEEIAFLKKVHEE EIRELQAQLQEQQVQVEMDMSKPDLTAALRDIRAQ
YETIAAKNISEAEEWYKSKVSDLTQAANKNNDALR QAKQEMMEYRHQIQSYTCEI-
DALKGTNDSLMRQMR ELEDRFASEASGYQDNIARLEEEIRHLKDEMARHL REYQDLL-
NVKMALDVEIATYRKLLEGEESRINLPI QTYSALNFRETSPEQRGSEVHTKKTVMIKTI-
ETRD GEVVSEATQQQHEVL

sp|P20648|ATP4A_HUMAN

Potassium-transporting ATPase alpha chain 1 OS=Homo sapiens GN=ATP4A PE=2 SV=5

MGKAENYELYSVELGPGPGGDMAAKMSKKKKAGGG GGKRKEKLENMKKEMEIND-
HQLSVAELEQKYQTSA TKGLSASLAAELLLRDGPNALRPPRGTPEYVKFAR QLAGGLQ-
CLMWVAAAICLIAFAIQASEGDLTTDDN LYLAIALIAVVVVTGCFGYYQEFKSTNIIASFKNL
VPQQATVIRDGDKFQINADQLVVGDLVEMKGGDRV PADIRILAAQGCKVDNSSLT-
GESEPQTRSPECTHE SPLETRNIAFFSTMCLEGTVQGLVVNTGDRTIIGR IASLAS-
GVENEKTPIAIEIEHFVDIIAGLAILFGA TFFIVAMCIGYTFLRAMVFFMAIVVAYVPEGLLAT
VTVCLSLTAKRLASKNCVVKNLEAVETLGSTSVIC SDKTGTLTQNRMTVSHLWFDNHIH-
TADTTEDQSGQ TFDQSSETWRALCRVLTLCNRAAFKSGQDAVPVPK RIVIGDASETAL-
LKFSELTLGNAMGYRDRFPKVCE IPFNSTNKFQLSIHTLEDPRDPRHLLVMKGAPERV
LERCSSILIKGQELPLDEQWREAFQTAYLSLGGLG ERVLGFCQLYLNEKDYPPGYAFD-
VEAMNFPSSGLC FAGLVSMIDPPRATVPDAVLKCRTAGIRVIMVTGD HPITAKAIAASVGI-
ISEGSETVEDIAARLRVPVDQ VNRKDARACVINGMQLKDMDPSELVEALRTHPEMV
FARTSPQQKLVIVESCQRLGAIVAVTGDGVNDSPA LKKADIGVAMGIAGSDAAKNAAD-
MILLDDNFASIV TGVEQGRLIFDNLKKSIAYTLTKNIPELTPYLIYI TVSVPLPLGCI-
TILFIELCTDIFPSVSLAYEKAES DIMHLRPRNPKRDRLVNEPLAAYSYFQIGAIQSFA
GFTDYFTAMAQEGWFPLLCVGLRAQWEDHHLQDLQ DSYGQEWTFGQRLYQQYT-
CYTVFFISIEVCQIADV LIRKTRRLSAFQQGFFRNKILVIAIVFQVCIGCFL CYCPGMP-
NIFNFMPIRFQWWLVPLPYGILIFVYDE IRKLGVRCCPGSWWDQELYY

3 Digest the protein sequence database

Now that the protein sequence database of the species of interest has been imported, you need to transform the protein sequence database into a peptide database. This will allow for the conversion of the SRL to the desired species via the protein sequence database.

To obtain the database of peptides, you must perform an *in silico* digestion of the proteins. The digestion in dialects only models trypsin with zero missed cleavages. It will also remove all peptides that are below 5 and above 52 amino acids long, and all peptides that are not unique to the database. This does not mean that these peptides are entirely unique and I encourage people to BLAST the peptides to check for true uniqueness.

```
digest_human <- digest.fasta(human_proteome)
```

accession

name

sequence

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

MFHLR

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

TCAAK

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

LRPLTASQTVK

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

TFSQNRPAAAR

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

TFQQIR

sp|Q02218|ODO1_HUMAN

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

CYSAPVAAEPFLSGTSSNYVEEMYCAWLENPK

4 Import a spectral reference library (SRL)

Once you have your database of peptides, you need to import your SRL so that you can perform the conversion. This package is currently compatable with PeakView, OneOmics and OpenSWATH SRL format. The PeakView and OneOmics SRLs must be tab separated txt files and OpenSWATH can be either comma or tab separated files (csv and tsv, respectively). Any other file format won't work.

Please note: At this time, this package cannot convert between a PeakView/OneOmics SRL and an OpenSWATH SRL.

4.1 PeakView and OneOmics

```
peakview_example <- system.file("extdata",
                                "rat_srl_example.txt",
                                package = "dialects")

rat_srl_pv <- import.srl(peakview_example,
                         SRL.format = "peakview")
```

Q1

Q3

RT_detected

protein_name

isotype

relative_intensity

stripped_sequence

modification_sequence

prec_z

frg_type

frg_z

frg_nr

iRT

uniprot_id

score

decoy

prec_y

confidence

shared

N

rank

mods

nterm

cterm

383.2107

187.0713

22.58767

Alpha-1-inhibitor 3 OS=Rattus norvegicus GN=A1i3 PE=1 SV=1

NA

1952.58

GEAFTLK

GEAFTLK

2

b

1

2

21.23

sp|P14046|A1I3_RAT

NA

False

13480.82

0.8023

False

109

NA

NA

NA

NA

428.2196

589.3304

10.79025

Aspartate aminotransferase, cytoplasmic OS=Rattus norvegicus GN=Got1 PE=1 SV=3

NA

255.00

EHDSVLR

EHDSVLR

2

y

1

5

8.54

sp|P13221|AATC_RAT

NA

False

1319.46

0.5735

False

120

NA

NA

NA

NA

490.7527

622.3307

27.46840

Aspartate aminotransferase, mitochondrial OS=Rattus norvegicus GN=Got2 PE=1 SV=2

NA

218663.00

MNLGVGAYR

MNLGVGAYR

2

y

1

6

26.48

sp|P00507|AATM_RAT
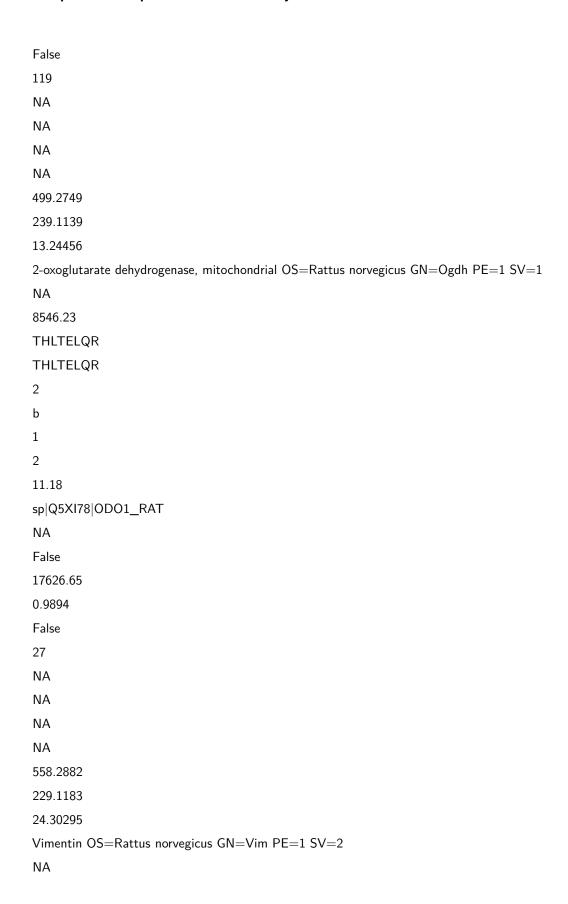
NA

False

359199.75

0.9980

False

77

NA

NA

NA

NA

370.7280

251.1503

16.62853

Sodium/potassium-transporting ATPase subunit alpha-1 OS=Rattus norvegicus GN=Atp1a1 PE=1 SV=1

NA

14240.77

HLLVMK

HLLVMK

2

b

1

2

14.82

sp|P06685|AT1A1_RAT

NA

False

21911.62

0.9954

False

93

NA

NA

NA

NA

426.7267

234.1448

18.06951

Sodium/potassium-transporting ATPase subunit alpha-2 OS=Rattus norvegicus GN=Atp1a2 PE=1 SV=1

NA

396070.60

YQVDLSK

YQVDLSK

2

y

1

2

16.37

sp|P06686|AT1A2_RAT

NA

False

588575.94

0.5160

False

17

NA

NA

NA

NA

645.5775

591.3110

71.97704

MICOS complex subunit Mic60 (Fragment) OS=Rattus norvegicus GN=Immt PE=1 SV=1

NA

35151.53

GMSISDLAGTLSTDDLNALIAHAHR

GMSISDLAGTLSTDDLNALIAHAHR

4

y

1

5

70.59

sp|Q3KR86|MIC60_RAT

1.762343e+07

False

207811.42

0.9997

False

123

NA

NA

NA

NA

387.7182

548.2609

16.15440

Unconventional myosin-Va OS=Rattus norvegicus GN=Myo5a PE=1 SV=1

NA

25328.65

IIGANMR

IIGANMR

2

y

1

5

14.31

sp|Q9QYF3|MYO5A_RAT

NA

False

20427.67

0.9391

False

16

NA

NA

NA

NA

558.6456

1020.5320

54.52056

Neurochondrin OS=Rattus norvegicus GN=Ncdn PE=1 SV=2

NA

15765.04

ILGAWLAEETSSLRK

ILGAWLAEETSSLRK

3

y

1

9

54.90

sp|O35095|NCDN_RAT

9.472999e+15

False

147196.98

1.0000

False

119

NA

NA

NA

NA

499.2749

239.1139

13.24456

2-oxoglutarate dehydrogenase, mitochondrial OS=Rattus norvegicus GN=Ogdh PE=1 SV=1

NA

8546.23

THLTELQR

THLTELQR

2

b

1

2

11.18

sp|Q5XI78|ODO1_RAT

NA

False

17626.65

0.9894

False

27

NA

NA

NA

NA

558.2882

229.1183

24.30295

Vimentin OS=Rattus norvegicus GN=Vim PE=1 SV=2

NA

26897.01

VELQELNDR

VELQELNDR

2

b

1

2

23.07

sp|P31000|VIME_RAT

5.931367e+18

False

101680.36

0.9993

False

94

NA

NA

NA

NA

4.2   OpenSWATH

```
openswath_example <- system.file("extdata",
                                 "rat_srl_example_openswath.tsv",
                                 package = "dialects")

rat_srl_os <- import.srl(openswath_example,
                         SRL.format = "openswath")
```

PrecursorMz

ProductMz

LibraryIntensity

PeptideSequence

FullUniModPeptideName

PrecursorCharge

FragmentType

FragmentCharge

FragmentSeriesNumber

Tr_recalibrated

UniprotID

CE

GroupLabel

decoy

ProteinName

Annotation

transition_group_id

PeptideGroupLabel

transition_name

428.2196

589.3304

255.00

EHDSVLR

EHDSVLR

2

y

1

5

8.54

sp|P13221|AATC_RAT

-1

light

0

sp|P13221|AATC_RAT

y5

01_EHDSVLR_2

01_EHDSVLR_2

01_y5_1_EHDSVLR_2

383.2107

187.0713

1952.58

GEAFTLK

GEAFTLK

2

b

1

2

21.23

sp|P14046|A1I3_RAT

-1

light

0

sp|P14046|A1I3_RAT

b2

02_GEAFTLK_2

02_GEAFTLK_2

02_b2_1_GEAFTLK_2

645.5775

591.3110

35151.53

GMSISDLAGTLSTDDLNALIAHAHR

GMSISDLAGTLSTDDLNALIAHAHR

4

y

1

5

70.59

sp|Q3KR86|MIC60_RAT

-1

light

0

sp|Q3KR86|MIC60_RAT

y5

03_GMSISDLAGTLSTDDLNALIAHAHR_4

03_GMSISDLAGTLSTDDLNALIAHAHR_4

03_y5_1_GMSISDLAGTLSTDDLNALIAHAHR_4

370.7280

251.1503

14240.77

HLLVMK

HLLVMK

2

b

1

2

14.82

sp|P06685|AT1A1_RAT

-1

light

0

sp|P06685|AT1A1_RAT

b2

04_HLLVMK_2

04_HLLVMK_2

04_b2_1_HLLVMK_2

387.7182

548.2609

25328.65

IIGANMR

IIGANMR

2

y

1

5

14.31

sp|Q9QYF3|MYO5A_RAT

-1

light

0

sp|Q9QYF3|MYO5A_RAT

y5

05_IIGANMR_2

05_IIGANMR_2

05_y5_1_IIGANMR_2

558.6456

1020.5320

15765.04

ILGAWLAEETSSLRK

ILGAWLAEETSSLRK

3

y

1

9

54.90

sp|O35095|NCDN_RAT

-1

light

0

sp|O35095|NCDN_RAT

y9

06_ILGAWLAEETSSLRK_3

06_ILGAWLAEETSSLRK_3

06_y9_1_ILGAWLAEETSSLRK_3

490.7527

622.3307

218663.00

MNLGVGAYR

MNLGVGAYR

2

y

1

6

26.48

sp|P00507|AATM_RAT

-1

light

0

sp|P00507|AATM_RAT

y6

07_MNLGVGAYR_2

07_MNLGVGAYR_2

07_y6_1_MNLGVGAYR_2

499.2749

239.1139

8546.23

THLTELQR

THLTELQR

2

b

1

2

11.18

sp|Q5XI78|ODO1_RAT

-1

light

0

sp|Q5XI78|ODO1_RAT

b2

08_THLTELQR_2

08_THLTELQR_2

08_b2_1_THLTELQR_2

558.2882

229.1183

26897.01

VELQELNDR

VELQELNDR

2

b

1

2

23.07

sp|P31000|VIME_RAT
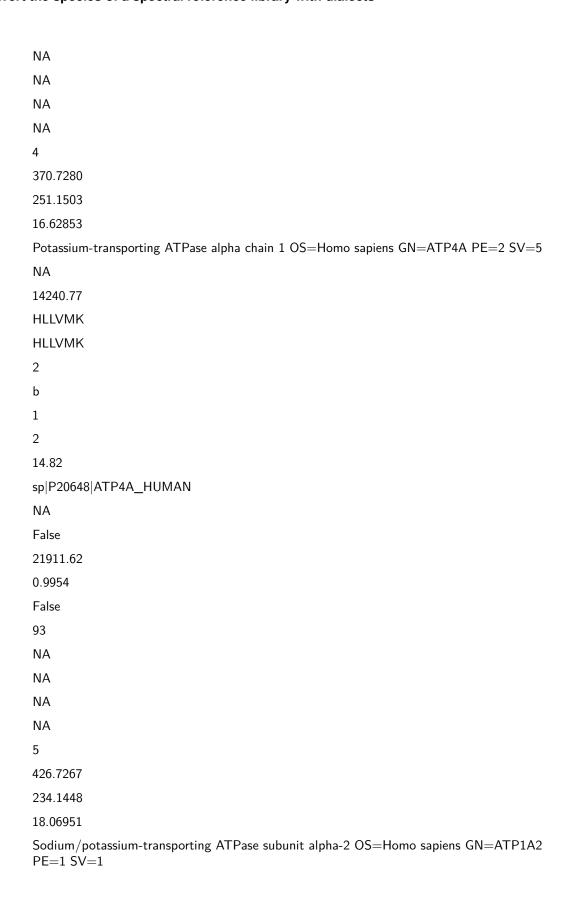
-1

light

0

sp|P31000|VIME_RAT

b2

09_VELQELNDR_2

09_VELQELNDR_2

09_b2_1_VELQELNDR_2

426.7267

234.1448

396070.60

YQVDLSK

YQVDLSK

2

y

1

2

16.37

sp|P06686|AT1A2_RAT

-1

light

0

sp|P06686|AT1A2_RAT

y2

10_YQVDLSK_2

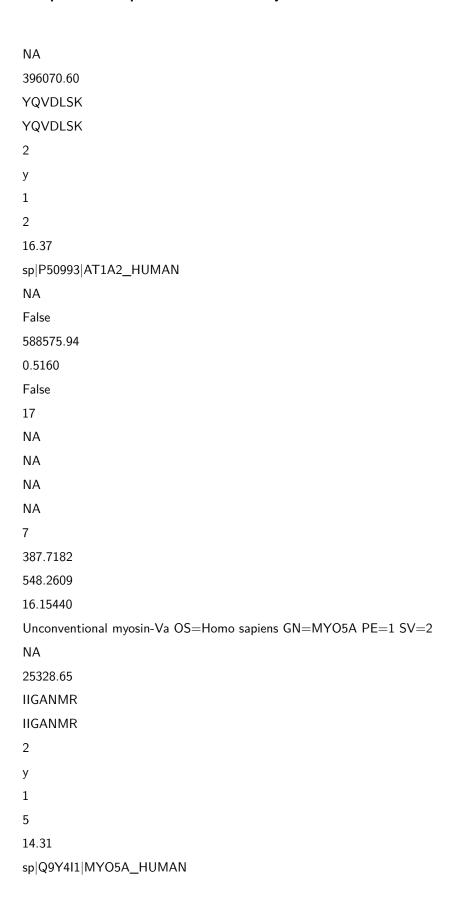10_YQVDLSK_2

10_y2_1_YQVDLSK_2
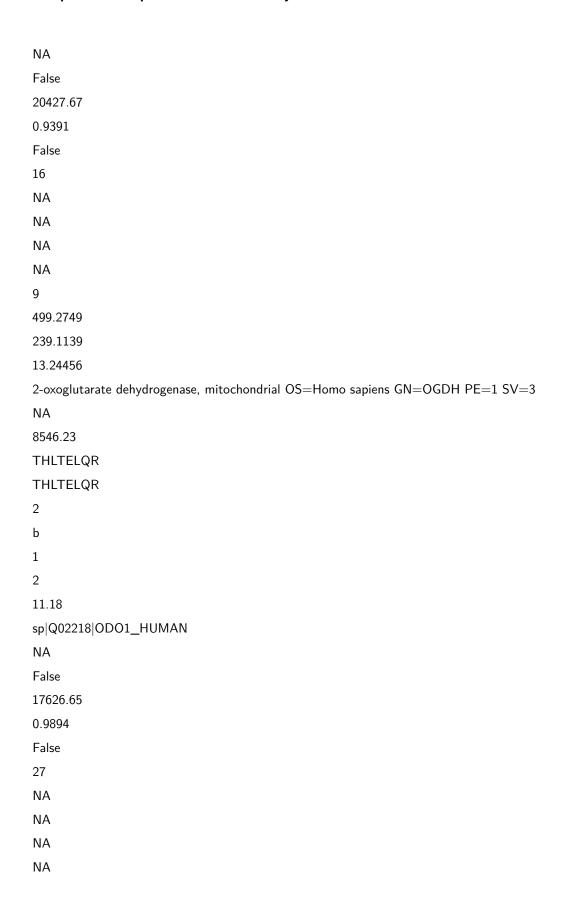
5   Covert the species of the SRL

Now you have the SRL and peptides from the protein sequence database, you can convert the species of your SRL. Please note: All retention time calibration peptides will not be converted and will not be copied to the new SRL.

```
human_from_rat <- convert.species(digest_human,
                                  rat_srl_pv,
                                  SRL.format = "peakview")
```

Convert the species of a spectral reference library with dialects

Q1

Q3

RT_detected

protein_name

isotype

relative_intensity

stripped_sequence

modification_sequence

prec_z

frg_type

frg_z

frg_nr

iRT

uniprot_id

score

decoy

prec_y

confidence

shared

N

rank

mods

nterm

cterm

1

383.2107

187.0713

22.58767

Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3

NA

1952.58

GEAFTLK

GEAFTLK

2

b

1

2

21.23

sp|P01023|A2MG_HUMAN

NA

False

13480.82

0.8023

False

109

NA

NA

NA

NA

3

490.7527

622.3307

27.46840

Aspartate aminotransferase, mitochondrial OS=Homo sapiens GN=GOT2 PE=1 SV=3

NA

218663.00

MNLGVGAYR

MNLGVGAYR

2

y

1

6

26.48

sp|P00505|AATM_HUMAN

NA

False

359199.75

0.9980

False

77

NA

NA

NA

NA

4

370.7280

251.1503

16.62853

Potassium-transporting ATPase alpha chain 1 OS=Homo sapiens GN=ATP4A PE=2 SV=5

NA

14240.77

HLLVMK

HLLVMK

2

b

1

2

14.82

sp|P20648|ATP4A_HUMAN

NA

False

21911.62

0.9954

False

93

NA

NA

NA

NA

5

426.7267

234.1448

18.06951

Sodium/potassium-transporting ATPase subunit alpha-2 OS=Homo sapiens GN=ATP1A2 PE=1 SV=1

NA

396070.60

YQVDLSK

YQVDLSK

2

y

1

2

16.37

sp|P50993|AT1A2_HUMAN

NA

False

588575.94

0.5160

False

17

NA

NA

NA

NA

7

387.7182

548.2609

16.15440

Unconventional myosin-Va OS=Homo sapiens GN=MYO5A PE=1 SV=2

NA

25328.65

IIGANMR

IIGANMR

2

y

1

5

14.31

sp|Q9Y4I1|MYO5A_HUMAN

NA

False

20427.67

0.9391

False

16

NA

NA

NA

NA

9

499.2749

239.1139

13.24456

2-oxoglutarate dehydrogenase, mitochondrial OS=Homo sapiens GN=OGDH PE=1 SV=3

NA

8546.23

THLTELQR

THLTELQR

2

b

1

2

11.18

sp|Q02218|ODO1_HUMAN

NA

False

17626.65

0.9894

False

27

NA

NA

NA

NA

6   Export the coverted SRL

The final step of the conversion process is to save the converted SRL. Unfortunately, this step cannot convert the SRL, thus the file format out the SRL must match the file format of the output. PeakView and OneOmics libraries can only be saved as txt files. OpenSWATH SRLs may be saved as either tsv or csv files.

```r
savepath <- system.file("extdata",
                         "human_from_rat_srl.txt",
                         package = "dialects")

export.srl(human_from_rat, savepath)
## [1] "Output complete"
```

7   Session info

```r
sessionInfo()
## R version 3.4.3 (2017-11-30)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS High Sierra 10.13.3
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_AU.UTF-8/en_AU.UTF-8/en_AU.UTF-8/C/en_AU.UTF-8/en_AU.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] kableExtra_0.7.0  knitr_1.19        dialects_0.0.0.56 BiocStyle_2.6.1
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.15        compiler_3.4.3      pillar_1.1.0
##  [4] plyr_1.8.4          highr_0.6           tools_3.4.3
##  [7] digest_0.6.15       evaluate_0.10.1     tibble_1.4.2
## [10] viridisLite_0.3.0   pkgconfig_2.0.1     rlang_0.1.6
## [13] yaml_2.1.16         xfun_0.1            stringr_1.2.0
## [16] httr_1.3.1          xml2_1.2.0          hms_0.4.1
## [19] rprojroot_1.3-2     ade4_1.7-10         data.table_1.10.4-3
## [22] R6_2.2.2            rmarkdown_1.8       bookdown_0.6
## [25] readr_1.1.1         seqinr_3.4-5        magrittr_1.5
## [28] backports_1.1.2     scales_0.5.0        htmltools_0.3.6
## [31] MASS_7.3-48         rvest_0.3.2         splitstackshape_1.4.2
## [34] colorspace_1.3-2    stringi_1.1.6       munsell_0.4.3
```