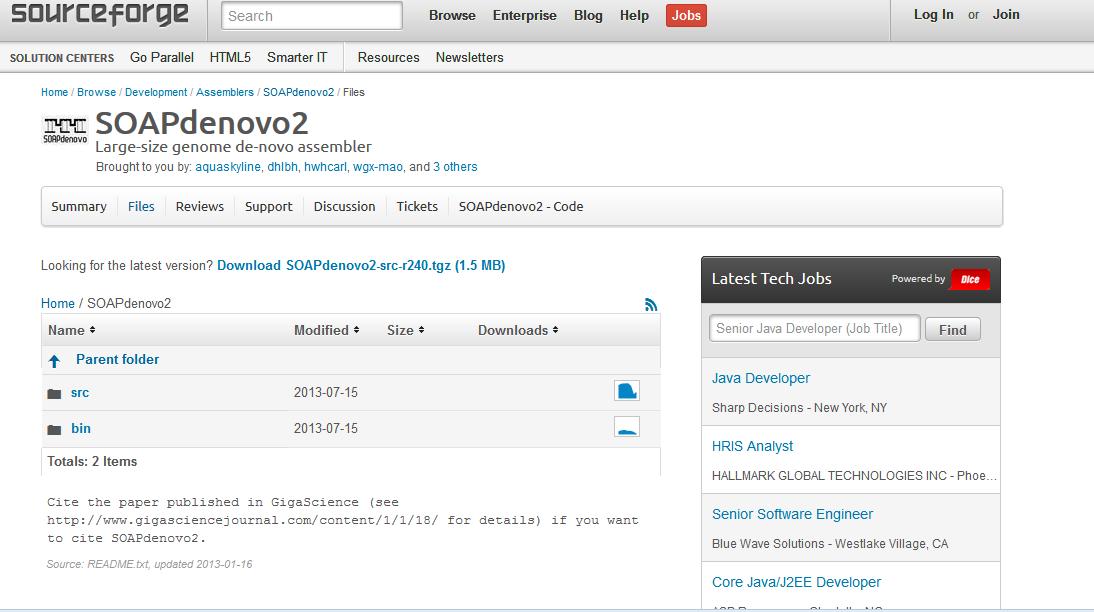
SOAPdenovo uitleg/userguide  
  
SOAP is an assembler tool which uses … and runs on all Linux distributions, it can be downloaded from … . SOAP can be found in two different flavours a bin and src version, the bin version only has to be extracted and is usable right after the src version instead has to be compiled for further use.  
SOAPdenovo can be downloaded from:

<http://sourceforge.net/projects/soapdenovo2/files/SOAPdenovo2/>



After one of the versions has been downloaded it can be extracted in the terminal with the command gunzip for the bin file and blabla for the src file.

When SOAP is compiled or extracted it can be started with …. but to actually run SOAP a configfile is needed, an example with commentary can be found below

with not before a config file is created, a template is created

make? alleen bij de scr versie

chmod +x SOAPdenovo-63mer

./SOAPdenovo-63mer

config file in dezelfde map als SOAPdenovo --

bin versie

config file:

#iets laten draaien

./SOAPdenovo-63mer all -s iets.config -K 63 -R -o graph\_prefix1 > ass.log 2 > ass.err

of 127mer, voor testen zie config.test in de soapdenovo map

q1=~/Downloads/C0A7AACXX\_101851-02\_TGACCA\_L001\_R2.fastq.gz

#maximal read length

max\_rd\_len=100

[LIB]

#average insert size

avg\_ins=200

#if sequence needs to be reversed

reverse\_seq=0

#in which part(s) the reads are used

asm\_flags=3

#use only first 100 bps of each read

rd\_len\_cutoff=100

#in which order the reads are used while scaffolding

rank=1

# cutoff of pair number for a reliable connection (at least 3 for short insert size)

pair\_num\_cutoff=3

#minimum aligned length to contigs for a reliable read location (at least 32 for short insert size)

map\_len=32

#weghalen wat niet wordt gebruikt

#maar vier files totaal of?

#a pair of fastq file, read 1 file should always be followed by read 2 file

q1=/path/\*\*LIBNAMEA\*\*/fastq1\_read\_1.fq

q2=/path/\*\*LIBNAMEA\*\*/fastq1\_read\_2.fq

#another pair of fastq file, read 1 file should always be followed by read 2 file

q1=/path/\*\*LIBNAMEA\*\*/fastq2\_read\_1.fq

q2=/path/\*\*LIBNAMEA\*\*/fastq2\_read\_2.fq

#a pair of fasta file, read 1 file should always be followed by read 2 file

f1=/path/\*\*LIBNAMEA\*\*/fasta1\_read\_1.fa

f2=/path/\*\*LIBNAMEA\*\*/fasta1\_read\_2.fa

#another pair of fasta file, read 1 file should always be followed by read 2 file

f1=/path/\*\*LIBNAMEA\*\*/fasta2\_read\_1.fa

f2=/path/\*\*LIBNAMEA\*\*/fasta2\_read\_2.fa

#fastq file for single reads

q=/path/\*\*LIBNAMEA\*\*/fastq1\_read\_single.fq

#another fastq file for single reads

q=/path/\*\*LIBNAMEA\*\*/fastq2\_read\_single.fq

#fasta file for single reads

f=/path/\*\*LIBNAMEA\*\*/fasta1\_read\_single.fa

#another fasta file for single reads

f=/path/\*\*LIBNAMEA\*\*/fasta2\_read\_single.fa

#a single fasta file for paired reads

p=/path/\*\*LIBNAMEA\*\*/pairs1\_in\_one\_file.fa

#another single fasta file for paired reads

p=/path/\*\*LIBNAMEA\*\*/pairs2\_in\_one\_file.fa

#bam file for single or paired reads, reads 1 in paired reads file should always be followed by reads 2

# NOTE: If a read in bam file fails platform/vendor quality checks(the flag field 0x0200 is set), itself and it's paired

read would be ignored.

b=/path/\*\*LIBNAMEA\*\*/reads1\_in\_file.bam

#another bam file for single or paired reads

b=/path/\*\*LIBNAMEA\*\*/reads2\_in\_file.bam

[LIB]

avg\_ins=2000

reverse\_seq=1

asm\_flags=2

rank=2

# cutoff of pair number for a reliable connection (at least 5 for large insert size)

pair\_num\_cutoff=5

#minimum aligned length to contigs for a reliable read location (at least 35 for large insert size)

map\_len=35

#weghalen wat niet wordt gebruikt

q1=/path/\*\*LIBNAMEB\*\*/fastq\_read\_1.fq

q2=/path/\*\*LIBNAMEB\*\*/fastq\_read\_2.fq

f1=/path/\*\*LIBNAMEA\*\*/fasta\_read\_1.fa

f2=/path/\*\*LIBNAMEA\*\*/fasta\_read\_2.fa

p=/path/\*\*LIBNAMEA\*\*/pairs\_in\_one\_file.fa

b=/path/\*\*LIBNAMEA\*\*/reads\_in\_file.bam

http://assemblathon.org/