

generateQuerySbjct, v. 0.4: Generate a Pair of Homologous DNA Sequences

Bernhard Haubold

Max-Planck-Institute for Evolutionary Biology, Plön, Germany

February 10, 2012

1 Introduction

2 Getting Started

generateQuerySbjct was written in C on a computer running Mac OS X and should work on any standard UNIX system. However, please contact me at haubold@evolbio.mpg.de if you have any problems with the program.

- Unpack the program

```
tar -xvzf generateQuerySbjct_XXX.tgz
```

where XXX indicates the version.

- Change into the newly created directory

```
cd GenerateQuerySbjct_XXX
```

and list its contents

```
ls
```

- Generate generateQuerySbjct

```
make
```

- List its options

```
./generateQuerySbjct -h
```

3 Listing

The following listing documents the driver program for generateQuerySbjct.

```
1 /***** generateQuerySbjct.c *****/
   * Description:
   * Author: Bernhard Haubold, haubold@evolbio.mpg.de
   * Date: Tue Sep 13 10:23:29 2011
   *****/
6 #include <stdio.h>
   #include <stdlib.h>
```

```

#include "interface.h"
#include "eprintf.h"

11 void generateSample(Args *args);

int main(int argc, char *argv[]){
    char *version;
    Args *args;

16    version = "0.4";
    setprogname2("generateQuerySbjct");
    args = getArgs(argc, argv);
    if(args->h || args->e)
21    printUsage(version);
    if(args->v)
        printSplash(version);
    generateSample(args);
    free(args);
26    free(progname());
    return 0;
}

void generateSample(Args *args){
31    FILE *pfp, *qfp, *sfp;
    char *templatel, *template2, *template3, *cmd;
    char *buf;
    int bufSize, numSeq;

36    bufSize = 1000;
    buf = (char *)emalloc(sizeof(char)*bufSize);

    templatel = "ms_2_1_s_%d_r_%f_%d_lms2dna_g_%f";
    template2 = "ms_2_1_t_%f_r_%f_%d_lms2dna_g_%f";
41    template3 = "macs_2_%d_t_%f_r_%f_2>/dev/null_lmsformatter_lms2dna_a_
        -g_%f";
    cmd = (char *)emalloc(sizeof(char)*bufSize);
    if(args->s)
        sprintf(cmd,templatel,args->s,args->r,args->l,args->g);
    else{
46        if(args->a){
            sprintf(cmd,template3,args->l,args->t/args->l,args->r/args->l,args->g
                );
        }else
            sprintf(cmd,template2,args->t,args->r,args->l,args->g);
    }
51    if((pfp = popen(cmd,"r")) == NULL){
        printf("ERROR[generateQuerySbjct]:_Could_not_open_pipe_for_command_'%s'
            '\n",cmd);
        exit(-1);
    }
    if((qfp = fopen(args->q,"w")) == NULL){
56    printf("ERROR[generateQuerySbjct]:_Could_not_open_'%s'_for_writing_query
        _sequence_to.\n",args->q);
        exit(-1);
    }

```

```

    }
    if((sfp = fopen(args->S, "w")) == NULL) {
        printf("ERROR[generateQuerySbjct]_Could_not_open_'%s'_for_writing_sbjct
            _sequence_to.\n", args->S);
61     exit(-1);
    }
    numSeq = 0;
    while(fgets(buf, 1000, pfp) != NULL) {
        if(buf[0] == '>')
66         numSeq++;
        if(numSeq == 1)
            fprintf(qfp, "%s", buf);
        else if(numSeq == 2)
            fprintf(sfp, "%s", buf);
71         else
            break;
    }
    free(buf);
    free(cmd);
76    fclose(qfp);
    fclose(sfp);
    pclose(pfp);
}

```

4 Change Log

- Version 0.3 (December 19, 2011)
 - Included option for using macs.
 - Improved user interface.
- Version 0.4 (December 20, 2011)
 - Included option for varying the GC-content.