

Objects

- ped: class Pedigree
 - ped.count: int, #individuals
 - ped.markerCount: int, #markers
 - ped.markerNames: StringArray, marker names
- ped[i]: class Person
 - ped[i].famid/pid/fatid/motid: String, pedigree info
 - ped[i].sex: int, sex
 - ped[i].traits: float*, traits
 - ped[i].markers: class Alleles*
 - ped[i].markers[j][0]/[1]: int, allele
 - ped[i].markers[j].one/two: int, allele, equivalent to item above
 - ped[i].father/mother: class Person*
 - ped[i].sibs: class Person**
 - ped[i].sibCount: int, sib count
- ped.GetMarkerInfo(j): class MarkerInfo
 - ped.GetMarkerInfo(j).CountAlleles(): int, #alleles
 - ped.GetMarkerInfo(j).freq[k]: float, allele frequency for allele k
 - ped.GetMarkerInfo(j).GetAlleleLabel(k): String, allele representation

Methods

- ped.EstimateFrequencies(1, true): method, calculate allele frequency based on founders

Example

CPP

```
#include <iostream>
#include "Core.hpp"
using namespace SEQLinco;

void showPed(Pedigree & ped)
{
    printf("Loaded %d individuals\n", ped.count);
    for (int i = 0; i < std::min(ped.count, 10); i++) {
        printf("[%s]: %s, %s, %s, %d\t\t",
            (const char *)ped[i].pid, (const char *)ped[i].famid,
            (const char *)ped[i].fatid, (const char *)ped[i].motid,
            ped[i].sex);
        for (int j = 0; j < ped.markerCount; ++j) {
            printf("%d%d\t", ped[i].markers[j].one, ped[i].markers[j].two);
        }
        printf("\n");
    }
    //
    printf("Loaded %d markers\n", ped.markerCount);
    // Estimate allele frequencies for all markers, verbose mode
    ped.EstimateFrequencies(1, false);
    // Get genotype statistics for markers
    for (int i = 0; i < ped.markerNames.Length(); i++) {
        printf("Statistics for marker [%s]\n", (const char *)ped.markerNames[i]);
    }
}
```

```

        // Allele index starts with 1 not 0
        for (int j = 1; j <= ped.GetMarkerInfo(i)->CountAlleles(); j++) {
            printf("\tFrequency for allele %d: %f\n", j, ped.GetMarkerInfo(i)->freq[j]);
            printf("\tName for allele %d: %s\n", j, (const char *)ped.GetMarkerInfo(i)->GetAlleleLabel(j));
        }
    }
    return;
}

void readData(Pedigree & ped,
              const char * datfile, const char * pedfile, const char * mapfile)
{
    // The data file contains a description of the contents of the
    // pedigree file, including for example, a list of marker and
    // trait names
    ped.Prepare(datfile);
    // The pedigree file contains a list of individuals, stored one
    // per row, with specific information about each individual as
    // detailed in the data file.
    ped.Load(pedfile);
    SortFamilies(ped);
    ped.LoadMarkerMap(mapfile);
    return;
}

void resetPed(Pedigree & ped)
{
    // I explore structure of PED by clearing it up ...
    // from PedigreeGlobal.h
    for (int i = 0; i < ped.markerInfoCount; i++) {
        ped.markerInfo[i]->freq.Clear();
        ped.markerInfo[i]->name.Clear();
        ped.markerInfo[i]->alleleLabels.Clear();
        ped.markerInfo[i]->alleleNumbers.Clear();
    }
    ped.markerNames.Clear();
    ped.markerLookup.Clear();
    ped.markerInfoByName.Clear();
    // from Pedigree.h
    for (int i = 0; i < ped.count; i++) {
        ped.persons[i]->famid.Clear();
        ped.persons[i]->pid.Clear();
        ped.persons[i]->motid.Clear();
        ped.persons[i]->fatid.Clear();
        ped.persons[i]->sex = ped.persons[i]->sibCount = ped.persons[i]->ngeno = ped.persons[i]->zygosity = 0;
    }
    for (int i = 0; i < ped.familyCount; i++) {
        ped.families[i]->famid.Clear();
        ped.families[i]->serial = ped.families[i]->first = ped.families[i]->last = ped.families[i]->count = ped.families[i]
    }->founders = ped.families[i]->nonFounders = ped.families[i]->generations = 0;
    }
    ped.pd.columns.Clear();
    ped.pd.columnHash.Clear();
    ped.pd.columnCount = 0;
    ped.markerCount = ped.markerInfoCount = ped.markerInfoSize = 0;
    ped.count = ped.familyCount = ped.haveTwins = 0;
    ped.size = 10000;
}

int main(int argc, char ** argv)
{
    if (argc != 3 && argc != 4) {
        printf("usage: %s <data source code: 1, 2, 3> <task: 1 or 2> -v\n", argv[0]);
        return 0;
    }

    Pedigree ped;
    std::string chrom = "1";
    if (atoi(argv[1]) == 1) readData(ped, "haplo.dat", "haplo.ped", "haplo.map");
    else if (atoi(argv[1]) == 2) readData(ped, "gene.dat", "gene.ped", "gene.map");
    else if (atoi(argv[1]) == 3) {
        std::vector<std::string> marker_ids { "V1", "V2", "V3" };
        std::vector<std::string> marker_positions { "1", "2", "3" };
        std::vector< std::vector<std::string> > samples;
    }
}

```

```

std::vector<std::string> s0 { "1", "1", "0", "0", "1", "21", "21", "21" };
samples.push_back(s0);
std::vector<std::string> s1 { "1", "2", "0", "0", "2", "11", "11", "11" };
samples.push_back(s1);
std::vector<std::string> s2 { "1", "3", "1", "2", "1", "21", "21", "21" };
samples.push_back(s2);
std::vector<std::string> s3 { "2", "1", "0", "0", "1", "22", "21", "00" };
samples.push_back(s3);
std::vector<std::string> s4 { "2", "2", "0", "0", "2", "11", "11", "11" };
samples.push_back(s4);
std::vector<std::string> s5 { "2", "3", "1", "2", "1", "21", "21", "21" };
samples.push_back(s5);
std::vector<std::string> s6 { "3", "1", "0", "0", "1", "22", "21", "21" };
samples.push_back(s6);
std::vector<std::string> s7 { "3", "2", "0", "0", "2", "11", "11", "21" };
samples.push_back(s7);
std::vector<std::string> s8 { "3", "3", "1", "2", "1", "21", "21", "21" };
samples.push_back(s8);
//
DataLoader dl;
dl.LoadVariants(ped, marker_ids, marker_positions, chrom);
dl.LoadSamples(ped, samples);
} else ;
if (atoi(argv[2]) == 1) showPed(ped);
else if (atoi(argv[2]) == 2) {
    MendelianErrorChecker mc;
    mc.Apply(ped);
    std::cout << "Mendelian Errors " << mc.errorCount << std::endl;
    GeneticHaplotyper gh(chrom);
    gh.Apply(ped);
    if (argc == 4) gh.Print();
    HaplotypeCoder hc(1);
    hc.Apply(gh.data);
    if (argc == 4) hc.Print();
} else ;
}

```

Makefile

```

CXX = g++
CFLAGS=-O3 -I ./ -I./libsrc -I./merlin -I./pdf -I./clusters -D_FILE_OFFSET_BITS=64 -D__ZLIB_AVAILABLE__ -Wall -std=c++11
BINDIR = demo
MERLIN = demo/demo.exe
EXECUTABLES = $(MERLIN)
# MERLIN File Set
MERLINBASE = merlin/AssociationAnalysis merlin/FastAssociation \
merlin/AnalysisTask merlin/Conquer \
merlin/ConquerHaplotyping merlin/DiseaseModel \
merlin/ParametricLikelihood merlin/GenotypeInference \
merlin/Houdini \
merlin/KongAndCox merlin/Manners merlin/MerlinBitSet \
merlin/MerlinCluster merlin/MerlinCore \
merlin/MerlinError merlin/MerlinFamily merlin/MerlinIBD \
merlin/InformationContent merlin/MerlinCache merlin/MerlinModel \
merlin/MerlinKinship merlin/MerlinKinship15 \
merlin/MerlinHaplotype merlin/MerlinMatrix merlin/MerlinParameters \
merlin/MerlinPDF merlin/MerlinSimulator merlin/MerlinSimwalk2 \
merlin/MerlinSort merlin/NPL-ASP merlin/NPL-QTL \
merlin/Magic merlin/Mantra merlin/Parametric merlin/QtlModel \
merlin/Tree \
merlin/TreeBasics merlin/TreeIndex merlin/TreeManager \
merlin/TreeInfo merlin/TreeFlips merlin/VarianceComponents
MERLINHDR = $(MERLINBASE:=.h) merlin/TreeNode.h Core.hpp
MERLINSRC = $(MERLINBASE:=.cpp) demo/LibmerlinDemo.cpp Core.cpp
MERLINOBJ = $(MERLINSRC:.cpp=.o)

# Files for dealing with clustered markers
CLUSTERS = clusters/HaploFamily clusters/HaploGraph \
clusters/HaploSet clusters/HaploTree clusters/Likelihood \
clusters/SparseLikelihood clusters/Unknown
CLUSTERCPP = $(CLUSTERS:=.cpp)
CLUSTERHDR = $(CLUSTERS:=.h)
CLUSTEROBJ = $(CLUSTERS:=.o)
# Utility Library File Set
LIBFILE = libsrc/lib-goncalo.a
LIBMAIN = libsrc/BasicHash libsrc/Error libsrc/FortranFormat \

```

```

libsrc/GenotypeLists libsrc/InputFile libsrc/IntArray libsrc/Hash \
libsrc/LongArray libsrc/Kinship libsrc/KinshipX libsrc/MapFunction \
libsrc/MathCholesky libsrc/MathDeriv libsrc/MathFloatVector \
libsrc/MathGenMin libsrc/MathGold libsrc/MathMatrix libsrc/MathStats \
libsrc/MathNormal libsrc/MathSVD libsrc/MathVector \
libsrc/MemoryInfo libsrc/MiniDeflate \
libsrc/Parameters libsrc/Pedigree libsrc/PedigreeAlleleFreq \
libsrc/PedigreeDescription libsrc/PedigreeFamily libsrc/PedigreeGlobals \
libsrc/PedigreePerson libsrc/QuickIndex libsrc/Random libsrc/Sort \
libsrc/StringArray libsrc/StringBasics libsrc/StringMap \
libsrc/StringHash libsrc/TraitTransformations
LIBPED = libsrc/PedigreeLoader libsrc/PedigreeTwin libsrc/PedigreeTrim
LIBSRC = $(LIBMAIN:=.cpp) $(LIBPED:=.cpp)
LIBHDR = $(LIBMAIN:=.h) libsrc/Constant.h \
libsrc/MathConstant.h libsrc/PedigreeAlleles.h libsrc/LongInt.h
LIBOBJ = $(LIBSRC:.cpp=.o)
# PDF Library File Sets
PDFLIB = pdf/libpdf.a
PDFFILES = pdf/PDF pdf/PDFfont pdf/PDFinfo pdf/PDFpage \
pdf/PDFchartbasics pdf/PDFchartbar pdf/PDFlinechart \
pdf/PDFhistogram \
pdf/PDFchartaxis pdf/PDFchartlegend pdf/PDFchartmarker \
pdf/PDFchartline pdf/PDFchartobject
PDFSRC = $(PDFFILES:=.cpp)
PDFHDR = $(PDFFILES:=.h)
PDFOBJ = $(PDFFILES:=.o)

# make everything
all : $(EXECUTABLES)

$(EXECUTABLES) : $(BINDIR)

$(BINDIR) :
    mkdir -p $(BINDIR)

# dependencies for executables
$(MERLIN) : $(LIBFILE) $(PDFLIB) $(MERLINOBJ) $(CLUSTEROBJ)
    $(CXX) $(CFLAGS) -o $@ $(MERLINOBJ) $(CLUSTEROBJ) $(PDFLIB) $(LIBFILE) -lm -lz

$(LIBFILE) : $(LIBOBJ) $(LIBHDR)
    ar -cr $@ $(LIBOBJ)
    ranlib $@

$(PDFLIB) : $(PDFOBJ)
    ar -cr $@ $(PDFOBJ)
    ranlib $@

$(MERLINOBJ) : $(MERLINHDR) $(CLUSTERHDR) $(LIBHDR)

$(CLUSTEROBJ) : $(CLUSTERHDR) $(MERLINHDR) $(LIBHDR)

$(LIBOBJ) : $(LIBHDR)

$(PDFOBJ) : $(PDFHDR)

clean :
    rm -f */*.a */*.o $(EXECUTABLES)

.c.o :
    $(CXX) $(CFLAGS) -o $@ -c $.c

#.cpp.X.o :
#    $(CXX) $(CFLAGS) -o $@ -c $.cpp -D__CHROMOSOME_X__

.cpp.o :
    $(CXX) $(CFLAGS) -o $@ -c $.cpp

.SUFFIXES : .cpp .c .o .X.o $(SUFFIXES)

```
