

ANALYSIS AND VISUALIZATION OF SSR DATA

USING R AND SHINY

SANMOHAN, JENS G. HANSEN, DAVID COOKE

U N E R S I T E T

ISOLATE STATISTICS

Under construction

Year 2014 2013 2012 2010 2009 2008 2007 2006 2005 2004 2002

Continent Europe Africa & West/Central Asia East Asia South America

Country All Belarus Belgium Bulgaria Croatia Czech Republic Denmark England Finland France Germany Greece Ireland Italy Netherlands Northern Ireland Norway Poland Portugal Scotland Slovenia Spain Sweden Switzerland Wales

[Input file](#)

SSR Data Analysis with Poppr

- [Data Summary](#) [Genotype Freq](#) [Min.Spanning Tree](#) [Phylogenetic Tree](#) [PCA](#) [MLG INFO](#) [Download](#) [Parallel Coord](#)

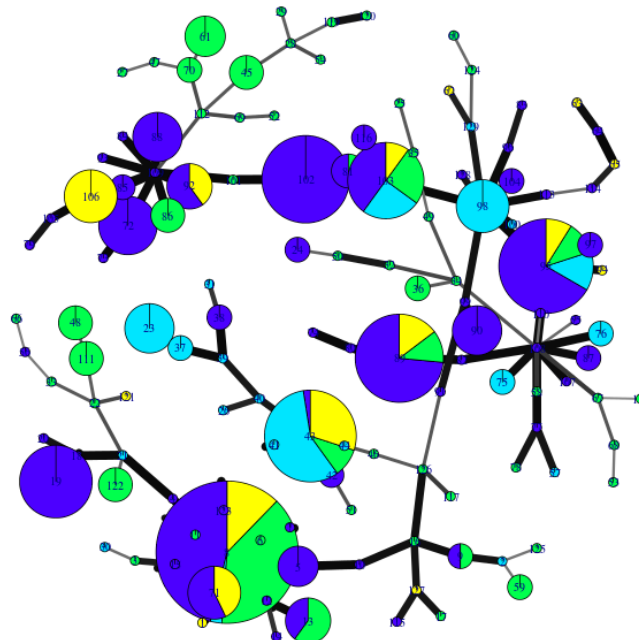
Label Nodes

Size MLG ID

Radio buttons can be used to set node labels of the tree. Select Size to get number of individuals in each MLG. Selecting MLG will display the MLG id of each node as its label. If ID is selected node labels will become individuals belonging to a node.

Choose a layout:

POPULATION



R PROGRAMMING

- › R is a free software programming language for statistical computing and graphics.
- › R's popularity has increased substantially in recent years
- › Poppr: An R package for analyzing population genetics data.

SHINY R

- › web framework for building interactive web applications using R.
- › computational power of R+interactivity of the modern web=shiny application

INTERACTIVE DATA ANALYSIS

- › Web interface to R computing using shiny.
- › Shiny makes analyses into interactive web applications that anyone can use.
- › Users choose input parameters using friendly controls like sliders, drop-downs, and text fields.
- › Easily incorporate any number of outputs like plots, tables, and summaries.

USE OF POPPR APPLICATION

- › Data overview
- › Data statistics
- › Compare populations and genotypes
- › ?
- › ?
- › ?
- › ?
- ›



ISOLATE STATISTICS

Under construction

Year 2014 2013 2012 2010 2009 2008 2007 2006 2005 2004 2002

Continent Europe Africa & West/Central Asia East Asia South America

Country All Belarus Belgium Bulgaria Croatia Czech Republic Denmark England Finland France Germany Greece Ireland Italy Netherlands Northern Ireland Norway Poland Portugal Scotland Slovenia Spain Sweden Switzerland Wales

Data selection

Input file

SSR Data Analysis with Poppr

Analysis options

Populations can be removed from further analysis. Activate below.

Activate subsetting

Yes No

Choose pop to be removed

Select Pop Structure for subsetting

genotypes Country Year

Collapse MLGs

Yes No

Choose a threshold below for collapsing MLGs. All MLGs less than threshold will be collapsed.

Threshold

0.05 0.5

Pop structure for analysis

genotypes Country Year

Refine data

No of multilocus genotypes:129
 No of Individuals:776
 No of codominant loci:12

	allele	1-D	Hexp	Evenness
Pi02	6.00	0.65	0.78	0.92
Pi04	3.00	0.52	0.78	0.92
Pi63	3.00	0.51	0.76	0.94
Pi70	2.00	0.42	0.84	0.86
D13	15.00	0.69	0.73	0.58
Pi4B	4.00	0.63	0.85	0.93
G11	15.00	0.56	0.60	0.48
SSR2	3.00	0.46	0.69	0.90
SSR4	9.00	0.76	0.85	0.88
SSR6	3.00	0.36	0.54	0.62
SSR8	3.00	0.50	0.76	0.97
SSR11	4.00	0.47	0.63	0.76
mean	5.83	0.54	0.73	0.81

Basic Table results

Pop	N	MLG	eMLG	SE	H	G	Hexp	E.5	Ia	rbarD
1 EN	377.00	57.00	20.37	2.74	2.36	4.11	0.76	0.33	4.23	0.47
2 ND	75.00	Data Hidden								
3 SC	235.00									
4 WA	89.00									
5 Total	776.00	129.00	24.62	3.26	2.66	3.91	0.75	0.22	4.00	0.42

POPPR APPLICATION IN ACTION

>

Server with
R and shiny

Results send back from shiny
server



Webpage

ISOLATE STATISTICS
Under construction

Year: 2014, 2013, 2012, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2002

Continent: Europe, Africa & West-Central Asia, East Asia, South America

Country: All, Belarus, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, England, Finland, France, Germany, Greece, Iceland, Italy, Netherlands, Northern Ireland, Norway, Poland, Portugal, Scotland, Slovenia, Spain, Sweden, Switzerland, Wales

SSR Data Analysis with Poppr

Data Summary | Genotype Freq | Min-Spanning Tree | Phylogenetic Tree | PCA | MLO BPO | Download | Parallel Coord

Choose populations to be removed: No of multiallelic genotypes: 24

Activate subsampling: No of individuals: 24

Choose pop to be removed: No of codominant loci: 12

allele	1-D	Hexp	Evenness
P102	3.00	0.23	0.34
P104	3.00	0.57	0.89
P163	3.00	0.46	0.98
P170	2.00	0.12	0.20
D13	11.00	0.74	0.40

Select Population Structure: Manual, Country, Year

User makes queries for
new
calculations

R sends data request

Data server



Data is returned

THANKS TO THE US GROUP – NIK
GRÜNWARD AND COLLEAGUES AT
OREGON STATE UNIVERSITY

THEY DEVELOPED AND PUBLISHED
THE CORE VERSION OF THIS TOOL