Melitaea britomartis population structure

Intermediate report

- Raw sequencing data QC
- Read mapping QC
- Variant calling summary
- Basic population structure (PCA)
- Mitochondrial tree (COI)

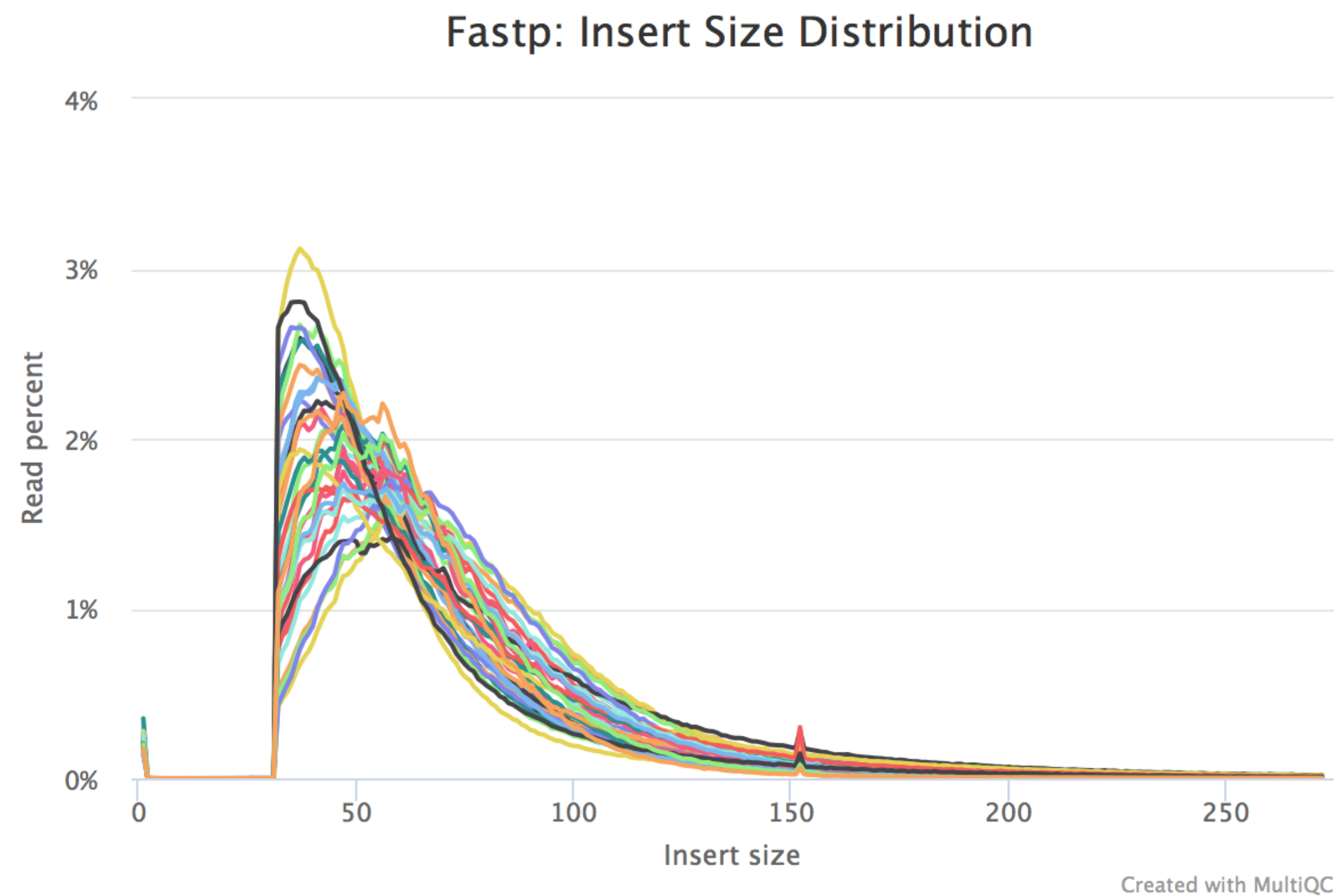# *Melitaea britomartis* population structure

Raw sequencing data QC

10-85M (single outlier 2.5M) of sequences reads obtained, most of the high quality
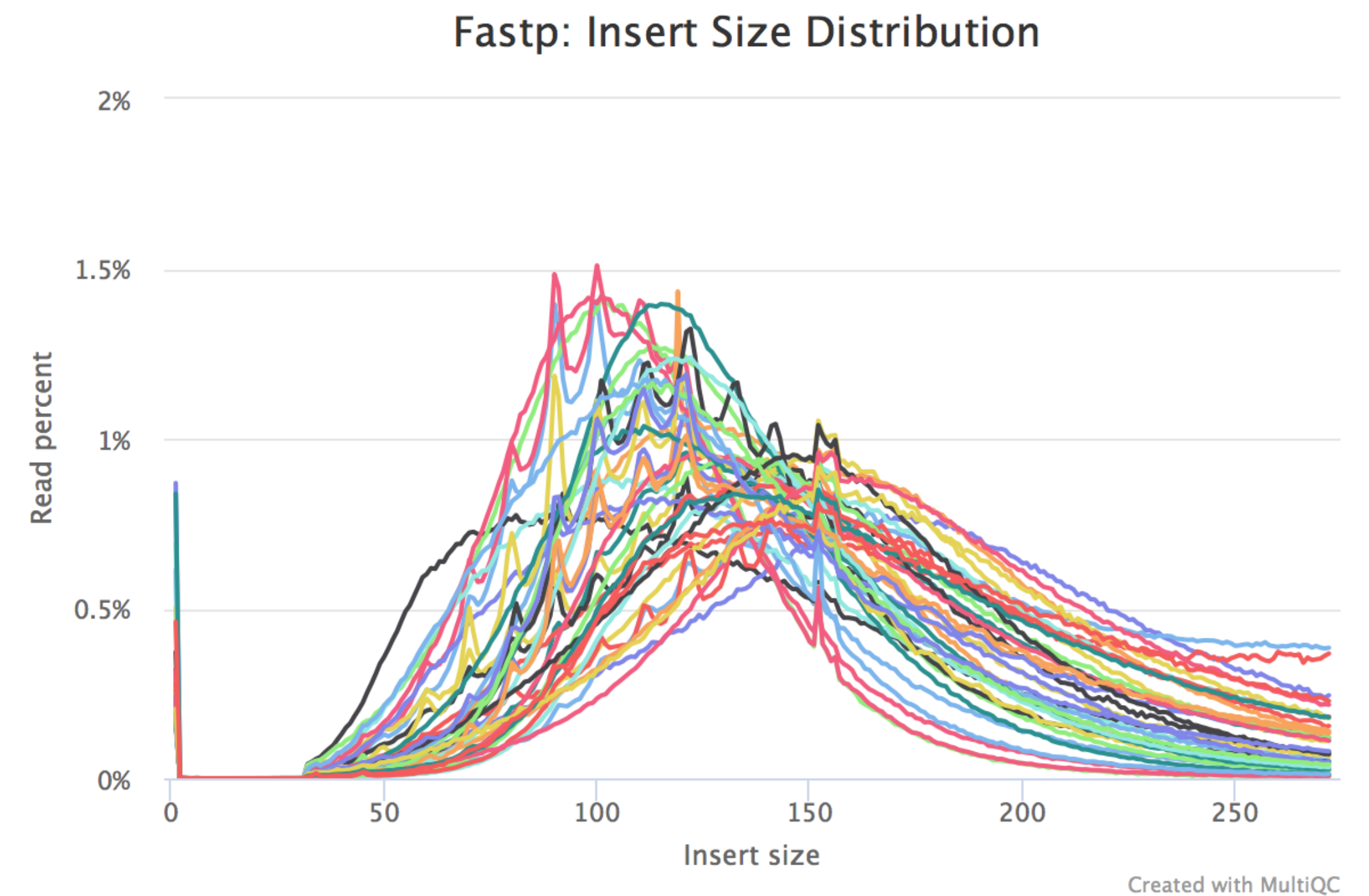GC content in large proportion of the samples appeared abnormal (37% is expected)
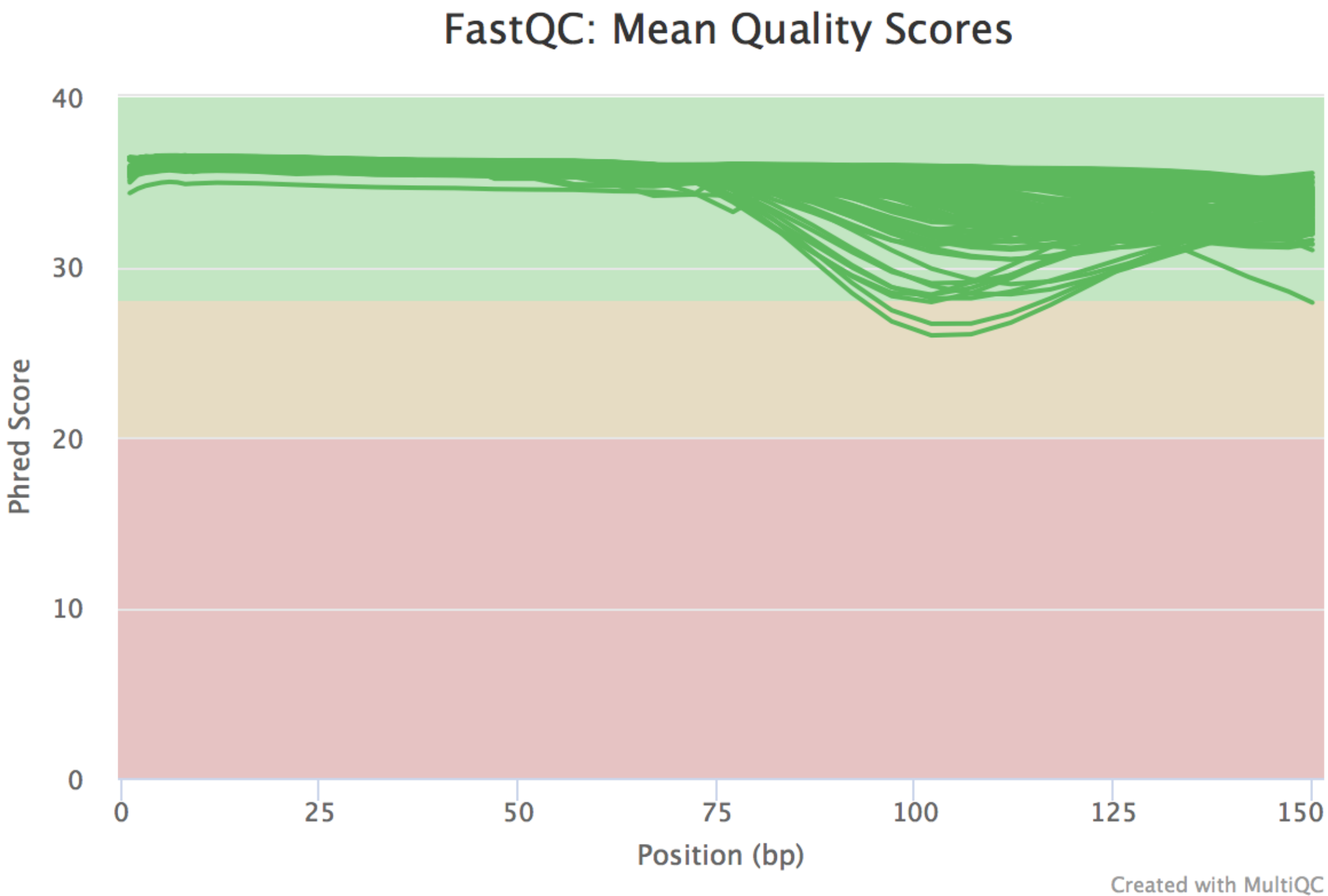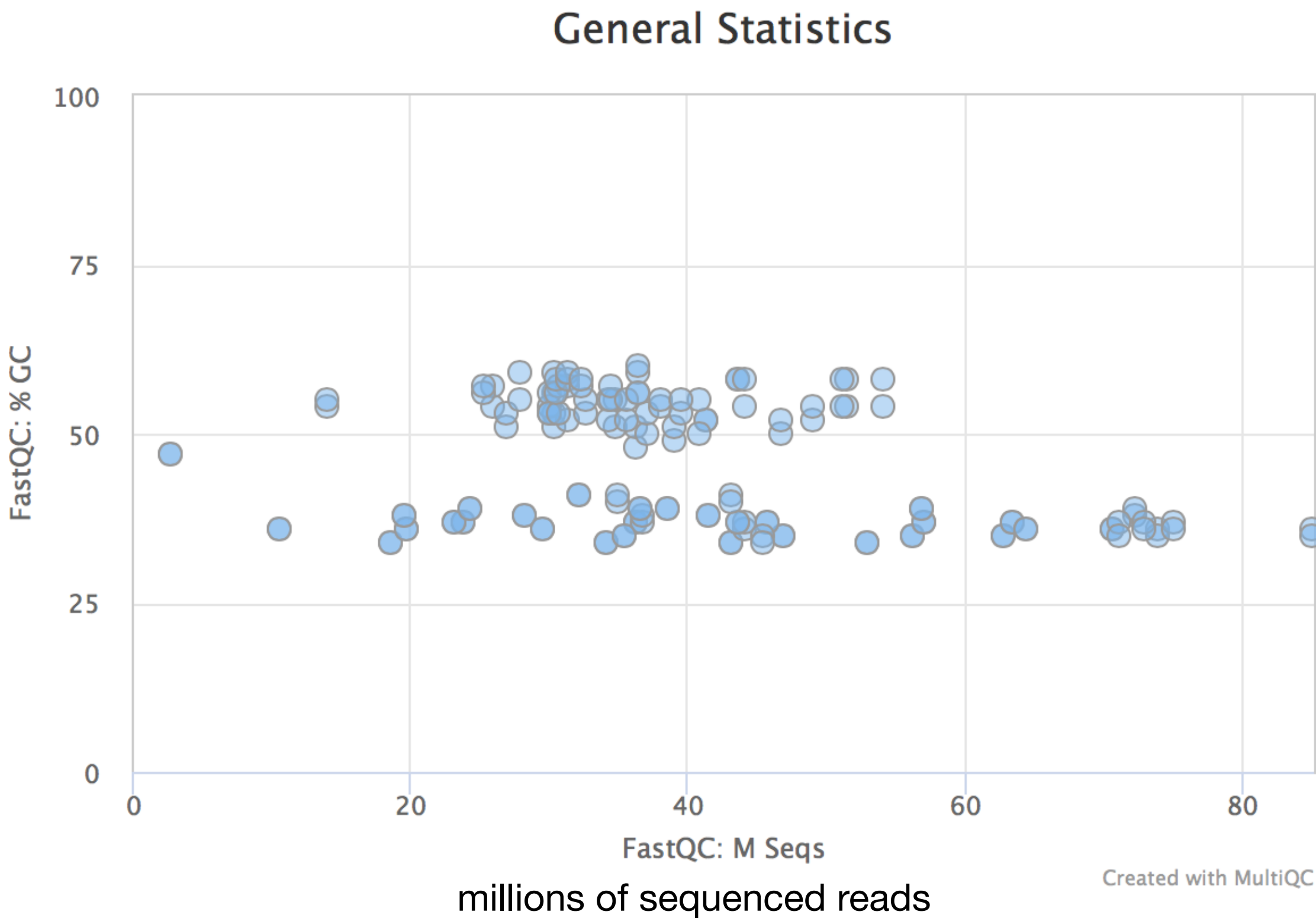
Melitaea britomartis population structure

Read mapping QC

Read mapping revealed difference in insert sizes (length of sequences fragments) between historical and contemporary samples
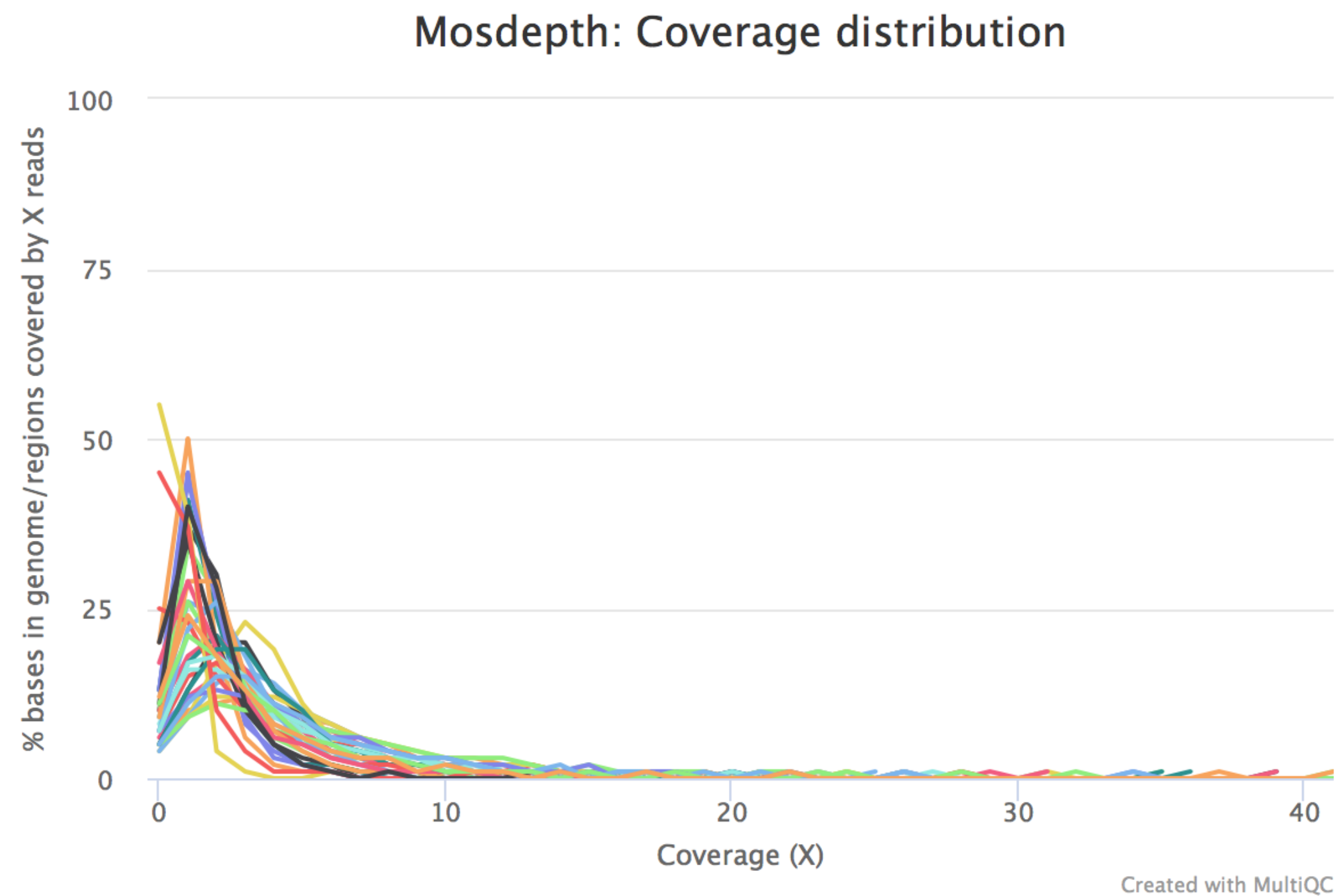

"Historical" samples
input GC > 50%

"Contemporary" samples
input GC < 50%

Melitaea britomartis population structure

Read mapping QC

Historical samples showed lower coverage / mapping depth, variant calling strategy adjusted to accommodate

"Historical" samples
input GC > 50%
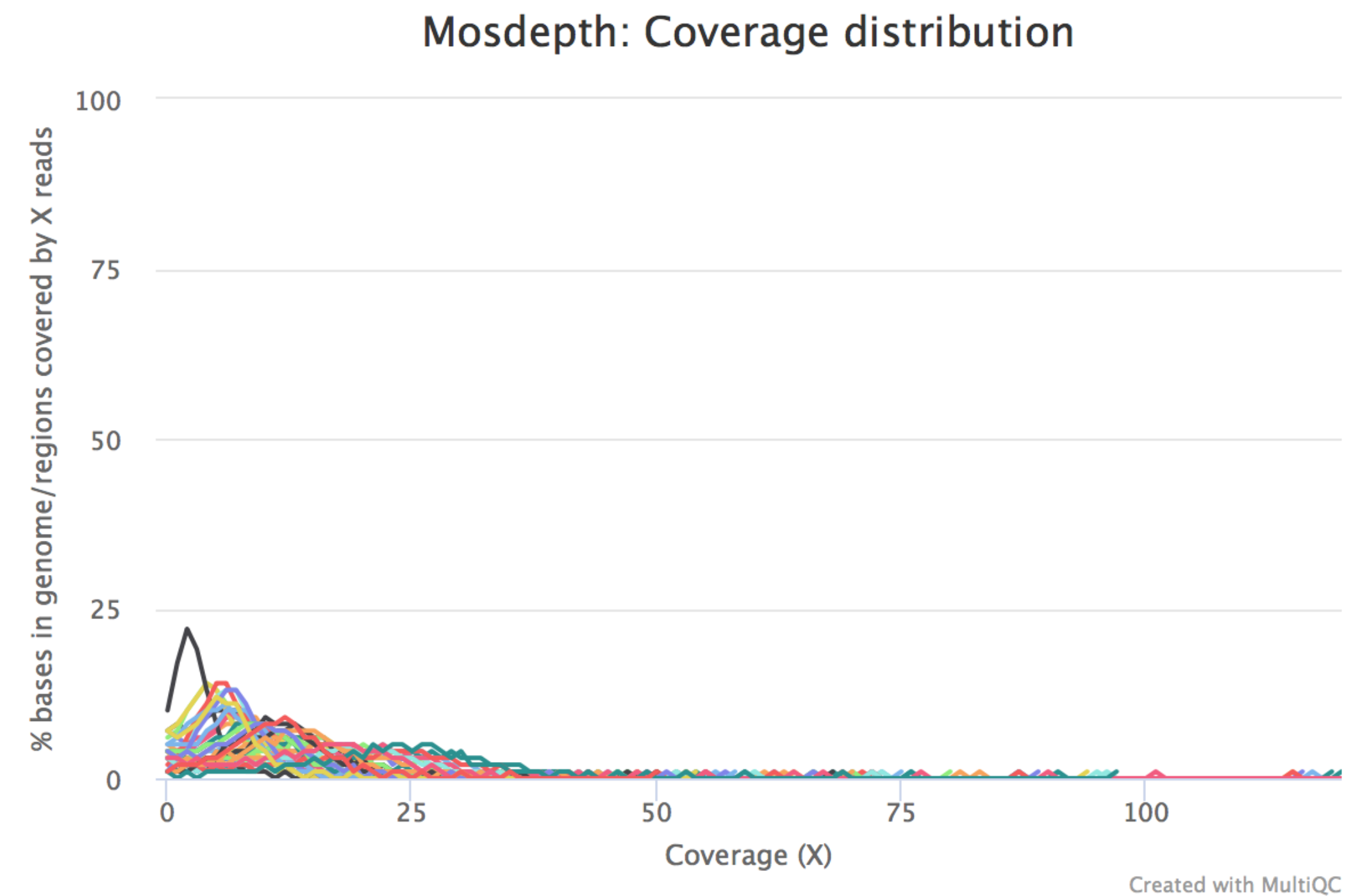
"Contemporary" samples
input GC < 50%

Melitaea britomartis population structure

Variant calling QC

At the first step samples are called individually (freeBayes), sufficient number of SNPs recovered
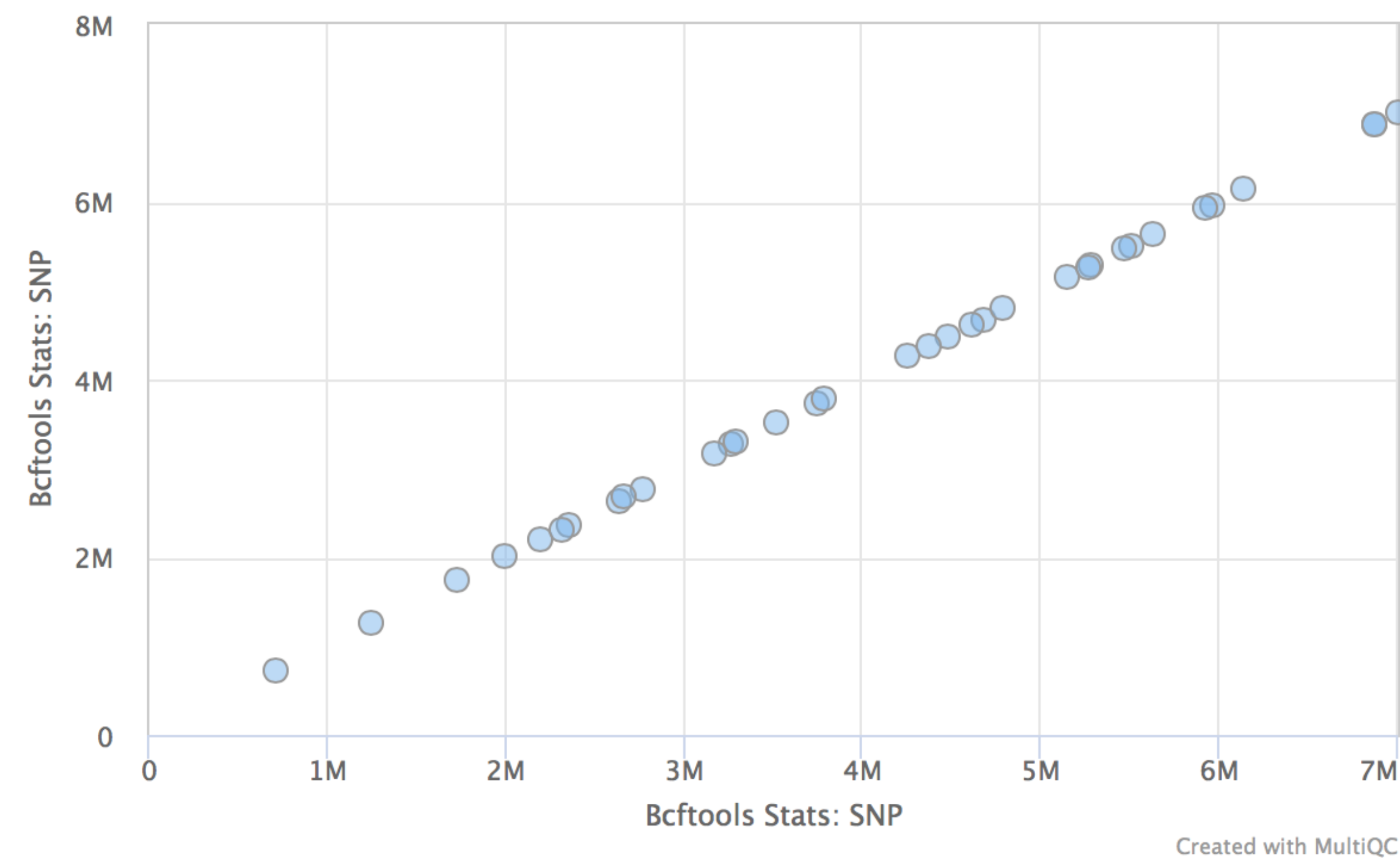
"Historical" samples
  input GC > 50%

"Contemporary" samples
  input GC < 50%

Melitaea britomartis population structure

Variant calling QC

Conclusion: distribution of substitution types doesn't indicate strong signatures of deamination
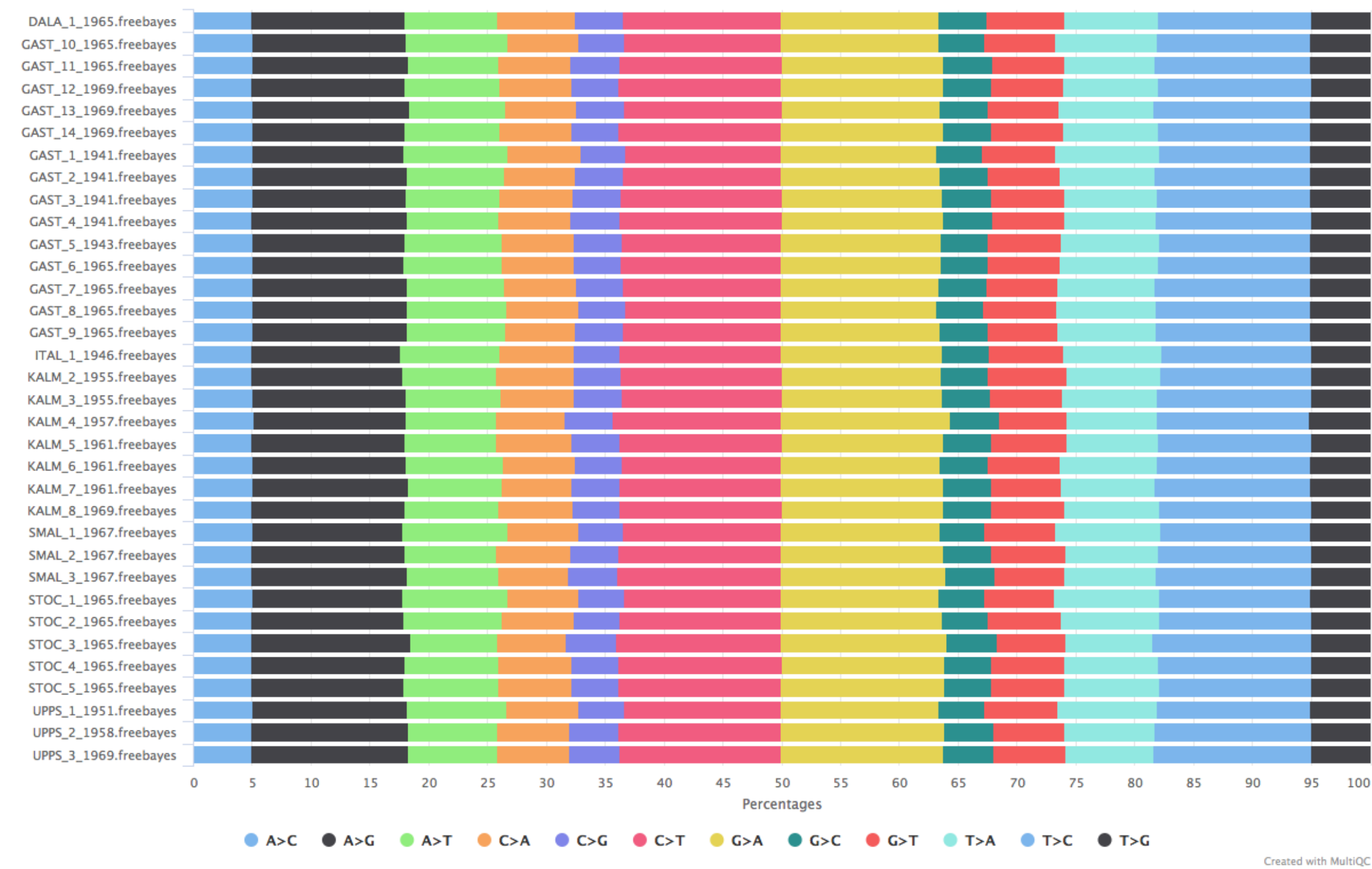
"Historical" samples
input GC > 50%
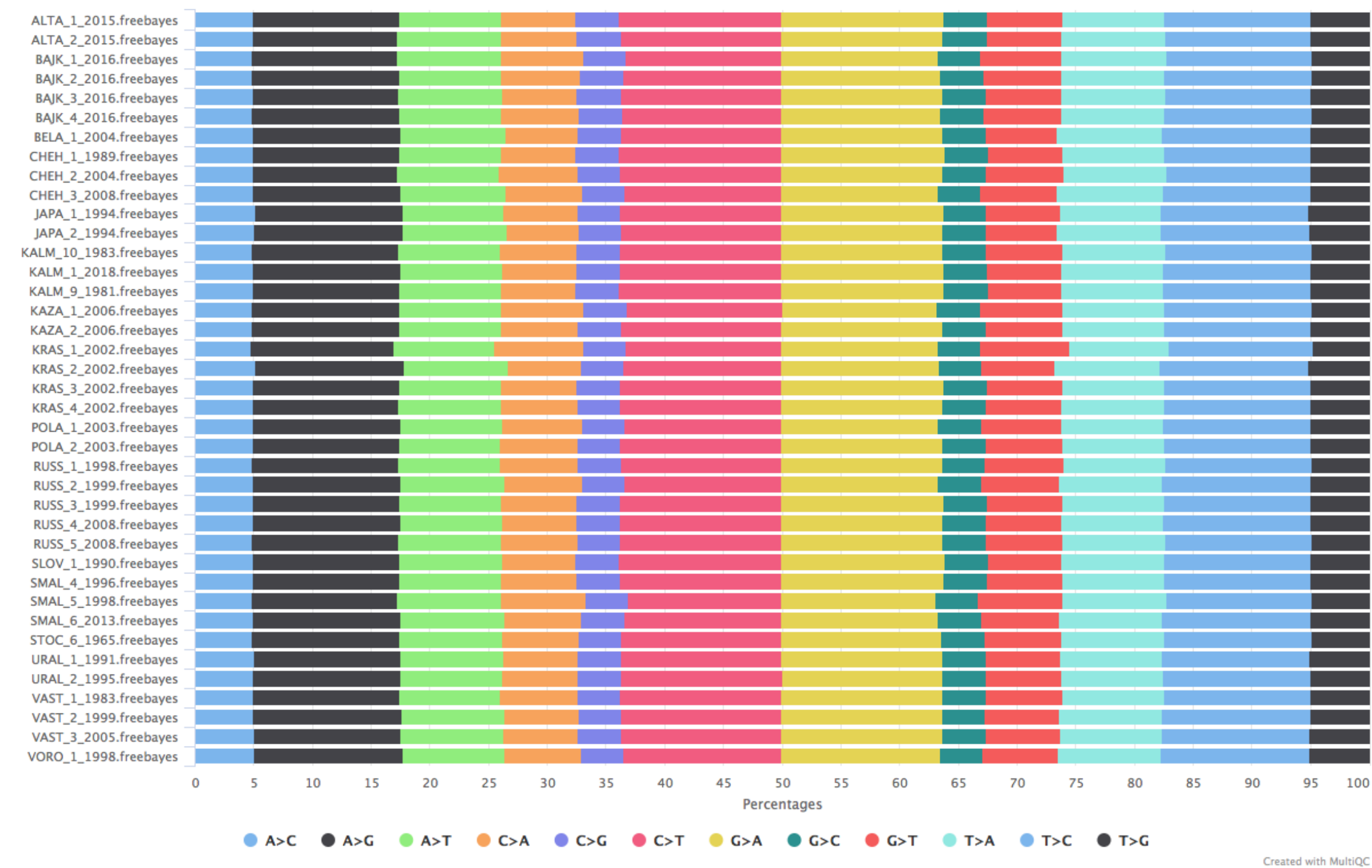
"Contemporary" samples
input GC < 50%

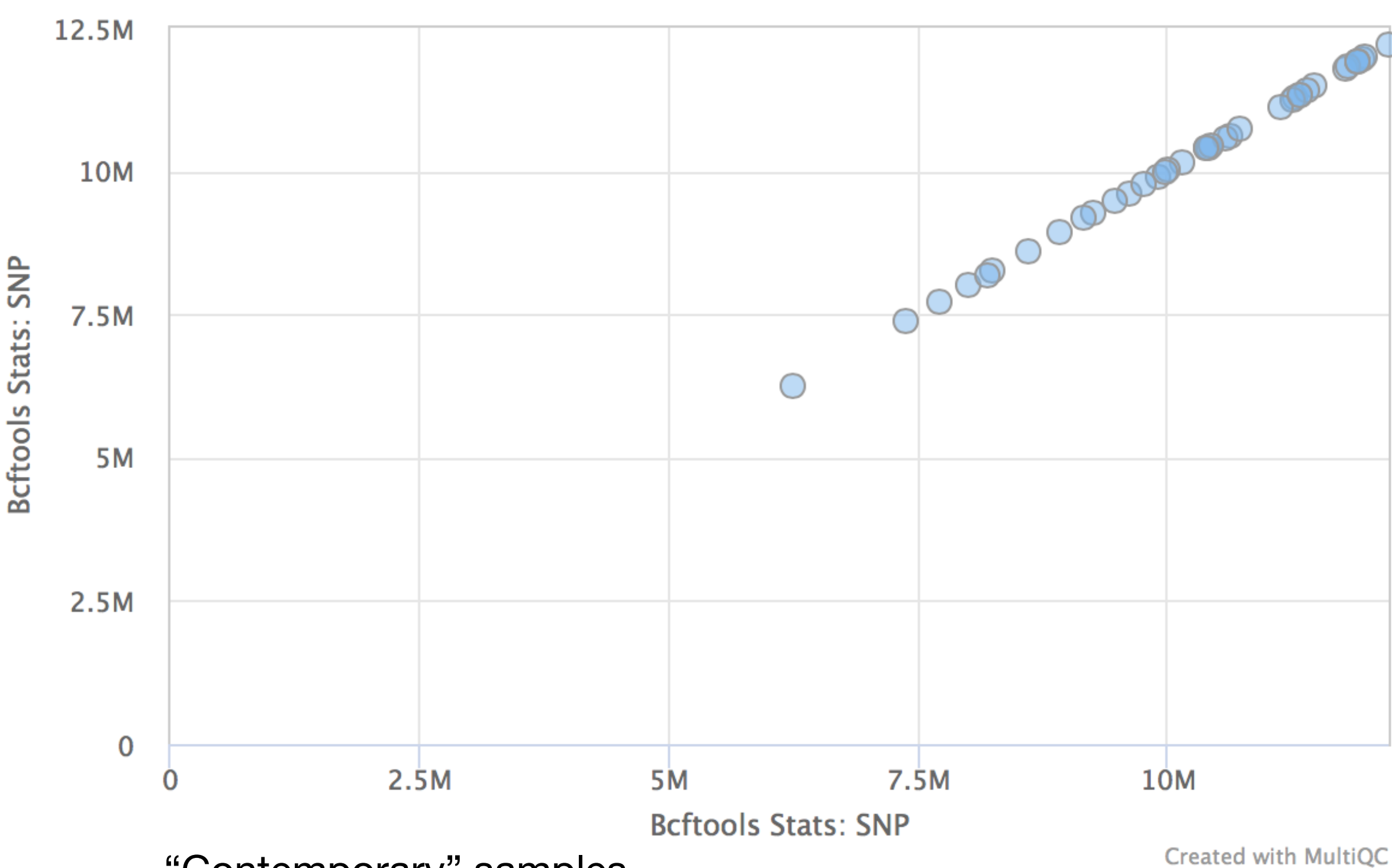# *Melitaea britomartis* population structure

Variant calling QC

Joined variant calling performed with settings from GenErode* pipeline

Stringient filtering applied:
        Quality filter: MQ < 30
        Missingness filter: variants are required to be present in at least 70 individuals

Total number of SNPs: **220,333**

*Kutschera, V.E., Kierczak, M., van der Valk, T. et al. GenErode: a bioinformatics pipeline to investigate genome erosion in endangered and extinct species. BMC Bioinformatics 23, 228 (2022). https://doi.org/10.1186/s12859-022-04757-0

Melitaea britomartis population structure

Population structure analysis

Basic principle component analysis (PCA) revealed unexpected grouping of samples based on geography and sampling year
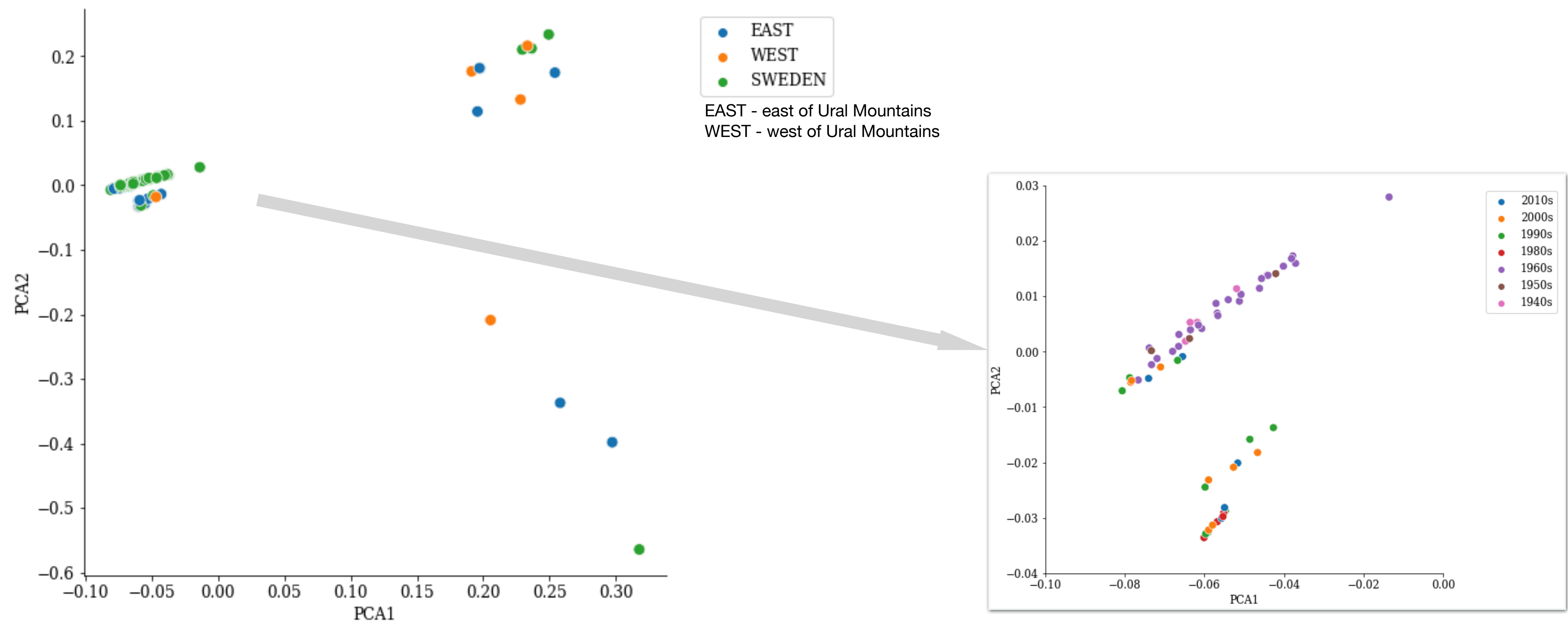
Melitaea britomartis population structure

Population structure analysis

Excluding mtDNA, sex-chromosomes and lower quality (historical) samples did not change overall structure of the plot

Melitaea britomartis

Mitochondrial tree (COI)

PCA outlier samples grouped with sister species

Kalmar 8 1969
Smaland 2 1967
Kalmar 3 1955
Kalmar 2 1955
Kalmar 1 2018
Japan 1 1994
Gastrikland 6 1965
Gastrikland 5 1943
Gastrikland 4 1941
Gastrikland 12 1969
Gastrikland 11 1965
Altai 1 2015
Krasnoyarsk 1 2002
Poland 2 2003
Russia 3 1999
Slovakia 1 1990
Stockholm 3 1965
URAL 2 1995
Stockholm 1 1965
Gastrikland 9 1965
Kalmar 6 1961
Gastrikland 13 1969
Stockholm 4 1965
Gastrikland 1 1941
Dalarna 1 1965
Uppsala 1 1951
Stockholm 2 1965
Gastrikland 10 1965
Gastrikland 2 1941
Gastrikland 7 1965
Gastrikland 8 1965
M.britomartis MW502384.1 cds QUA04452.1 1
Stockholm 5 1965
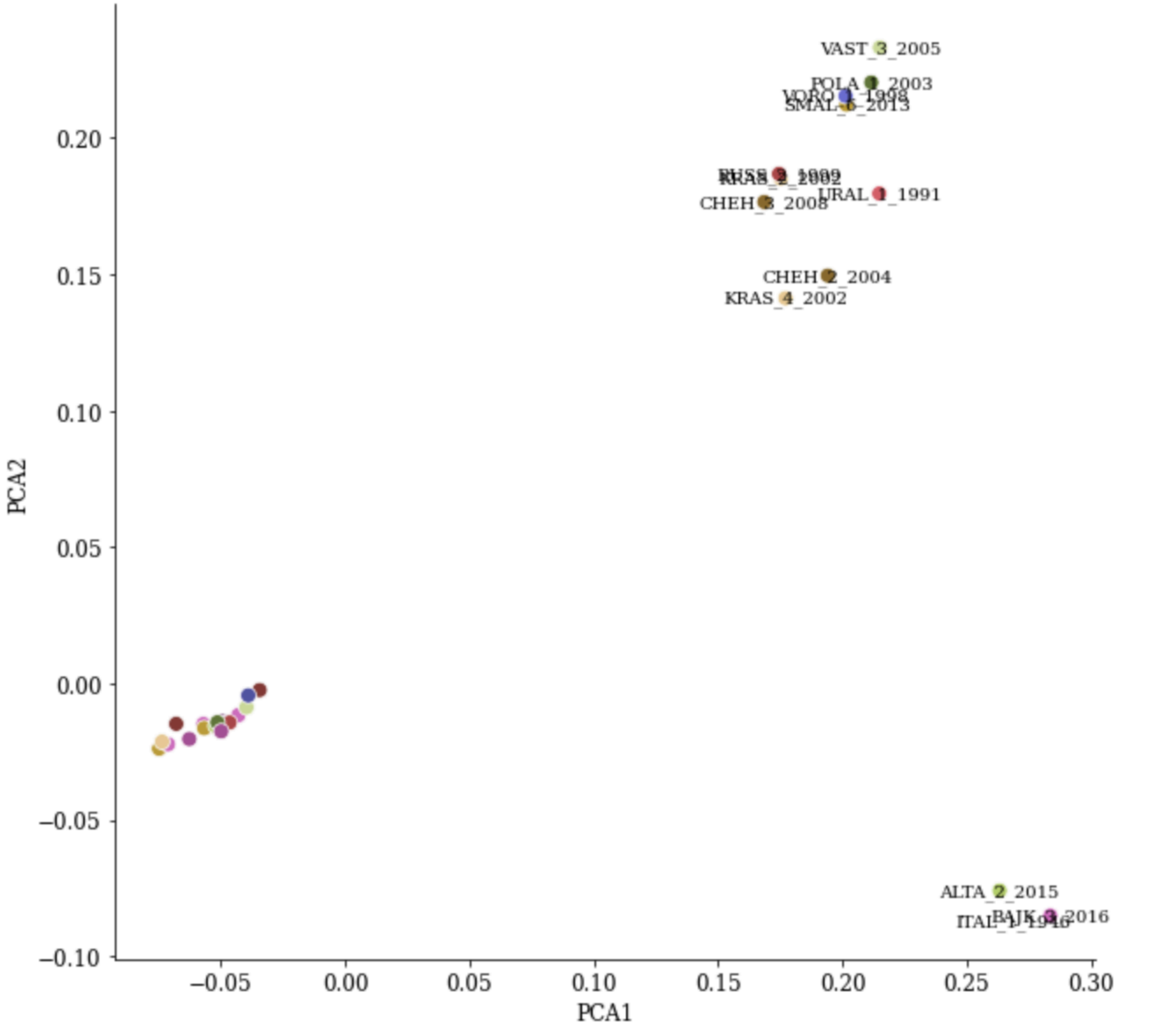Gastrikland 14 1969
Gastrikland 3 1941
Kalmar 5 1961
Kalmar 10 1983
Kalmar 9 1981
Kazakhstan 2 2006
Krasnoyarsk 3 2002
Vastmanland 1 1983
M.britomartis MW502388.1 cds QUA04456.1 1
Bajkal 1 2016
Bajkal 4 2016
Japan 2 1994
Kalmar 7 1961
Kazakhstan 1 2006
Russia 4 2008
Smaland 1 1967
Smaland 3 1967
Smaland 4 1996
Smaland 5 1998
Stockholm 6 1965
Uppsala 2 1958
Uppsala 3 1969
Russia 1 1998
Bajkal 2 2016
Russia 5 2008
Belarus 1 2004
Czechia 1 1989
M.britomartis MW502142.1 cds QUA04210.1 1
Vastmanland 2 1999
Czechia 2 2004
Krasnoyarsk 4 2002
Poland 1 2003
M.athalia HQ004748.1 cds ADM17493.1 1
Krasnoyarsk 2 2002
Ural 1 1991
M.athalia genome ref
Russia 2 1999
M.athalia HQ004750.1 cds ADM17495.1 1
M.athalia HQ004749.1 cds ADM17494.1 1
Voronezh 1 1998
Vastmanland 3 2005
Czechia 3 2008
Smaland 6 2013
Italy 1 1946
Altai 2 2015
Bajkal 3 2016
Kalmar 4 1957
M.diamina MW503608.1 cds QUA05676.1 1
M.diamina MW503503.1 cds QUA05571.1 1
M.cinxia MW501866.1 cds QUA03934.1 1

0.01