Landscape genomics

Physalia courses , November 26-30, 2018, Berlin

Spatial overlay & spatial coincidence concept

Dr Stéphane Joost – Oliver Selmoni (Msc)

Laboratory of Geographic Information Systems (LASIG)

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Spatial coincidence

- Landscape genomics mainly uses one of the six concepts of spatial analysis
- In order to connect genetic information with geo-environmental data, we utilize **spatial coincidence** analysis
- This approach associates information levels and is able to compare across them thanks to their common geographical coordinates
- We need a geo-referenced data set comprising one or more environmental variables describing the sampling location and a geo-referenced molecular marker data set for the study individuals or population(s)

PERSPECTIVES

Science, 2010

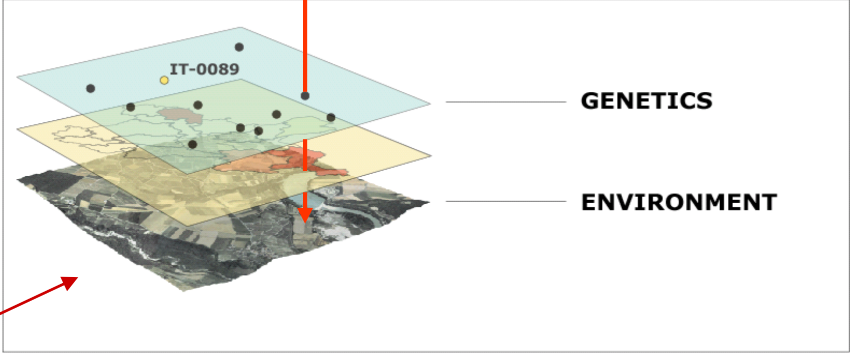
ECOLOGY

Time to Tap Africa's Livestock Genomes

Olivier Hanotte,[1] Tadelle Dessie,[2] Steve Kemp[3]

Fortunately, the fields of genetics and genomics (*3–5*) offer a new start for the sustainable improvement of African livestock productivity. Landscape genomics links genome-wide information to geo-environmental resource analysis to identify potentially valuable genetic material. Typically, researchers will perform a genome-wide scan on a number of animals from populations living in different habitats or across an ecological cline (from dry to wet areas, for instance).

Landscape genomics



- Correlative approaches and spatial statistics

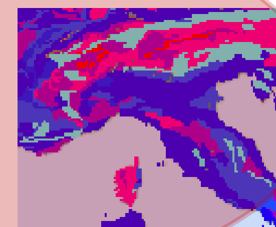
Genome-wide information

- Paradigm shift and transition phase

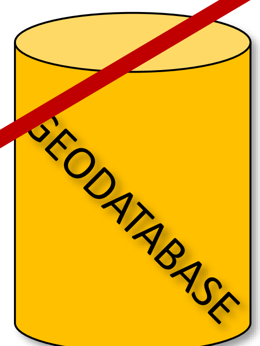
Data integration

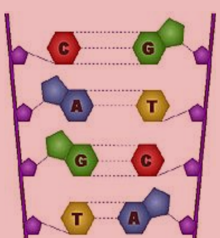
Administrative boundaries
Socio-Economic data
Socio-demographic data

Environmental data: topography, climate, soil, etc.



Geographic coordinates X, Y





Sampling
Genetic data

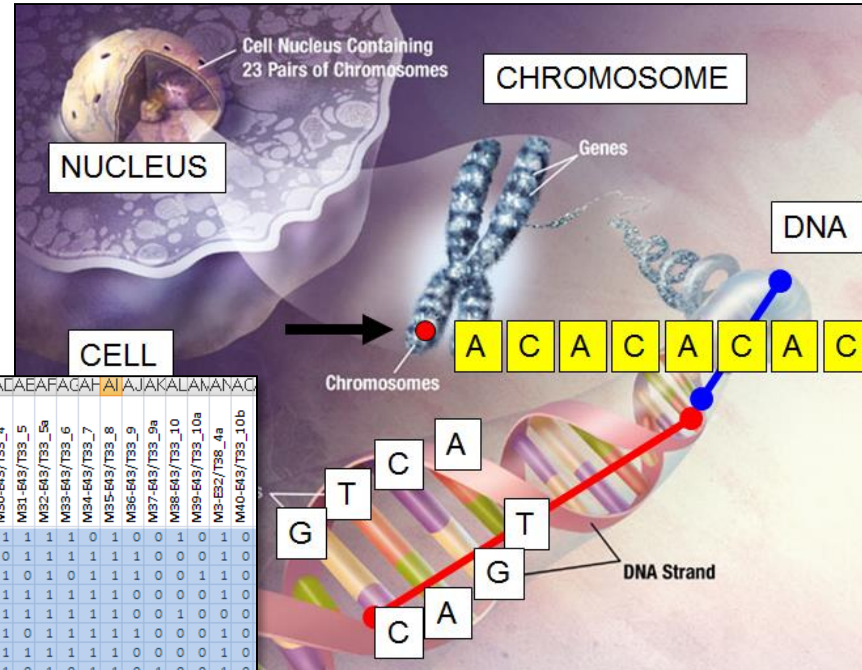
GIS

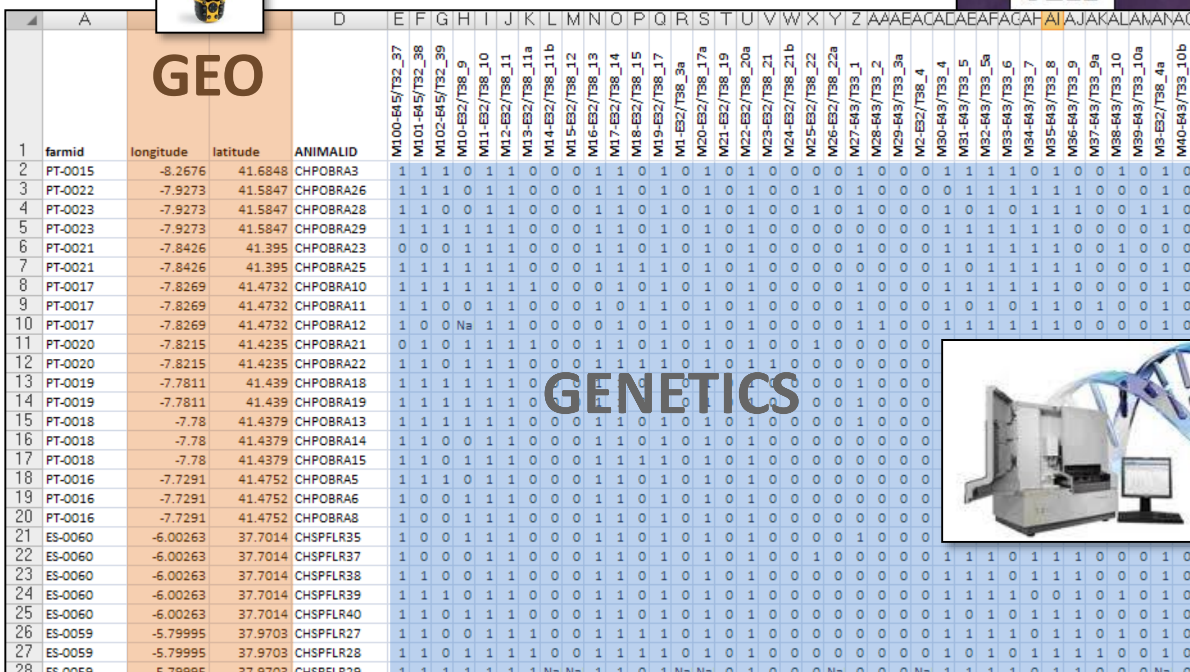
Questionnaires
Husbandry pratices

Transfer of attributes

- Usually we start with a file containing the coordinates of the sampling locations
- Then we add molecular information
- The next step is to feed this geo-molecular matrix with environmental descriptors

Complete the genetic matrix…





LECA Laboratoire d'Ecologie Alpine

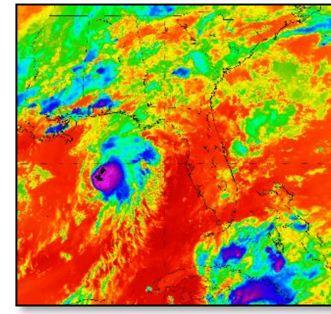
CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD School of Biosciences

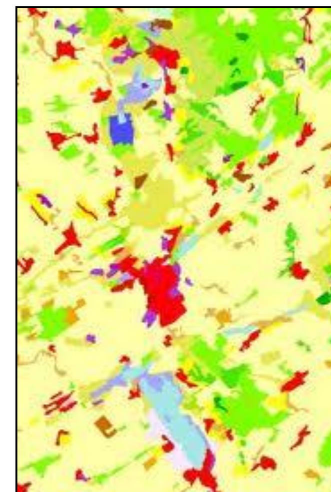
EMBL-EBI European Bioinformatics Institute

…with environmental variables characterizing sampling locations

GEO | GENETICS | ENVIRONMENT

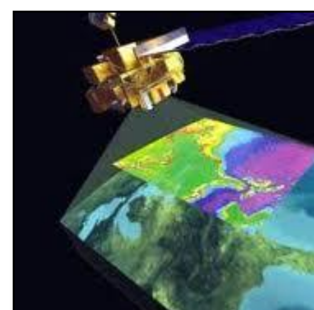
1	farmid	animalid	0ARJMP29_allele2_137	0ARJMP29_allele2_139	0ARJMP29_allele2_141	0ARJMP29_allele2_143	0ARJMP29_allele2_145	0ARJMP29_allele2_147	0ARJMP29_allele2_149	0ARJMP29_allele2_151	0ARJMP29_allele2_153	0ARJMP29_allele2_155	0ARJMP29_allele2_157	0ARJMP29_allele2_159	0ARJMP29_allele2_161	0ARJMP29_allele2_163	0ARJMP29_allele2_165	wndjan	altitude	wndfeb	wndmar	wndapr
1044	PL-4005	0APLPOM25	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.1	22	4.6	5	4.4
1045	PL-4005	0APLPOM26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.1	22	4.6	5	4.4
1046	PL-4006	0APLPOM01	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5.3	153	4.8	4.9	4.3
1047	PL-4006	0APLPOM15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.3	153	4.8	4.9	4.3
1048	PL-4006	0APLPOM24	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5.3	153	4.8	4.9	4.3
1049	PL-4007	0APLPOM05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.3	250	4.8	5	4.5
1050	PL-4007	0APLPOM16	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5.3	250	4.8	5	4.5
1051	PL-4008	0APLPOM09	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5.2	166	4.8	5	4.4
1052	PL-4008	0APLPOM19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.2	166	4.8	5	4.4
1053	PL-4008	0APLPOM20	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	5.2	166	4.8	5	4.4
1054	PL-4009	0APLPOM10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.5	87	5	5.2	4.6
1055	PL-4009	0APLPOM21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.5	87	5	5.2	4.6
1056	PL-4010	0APLPOM08	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5.4	208	4.9	5.1	4.5





esa

European Environment Agency

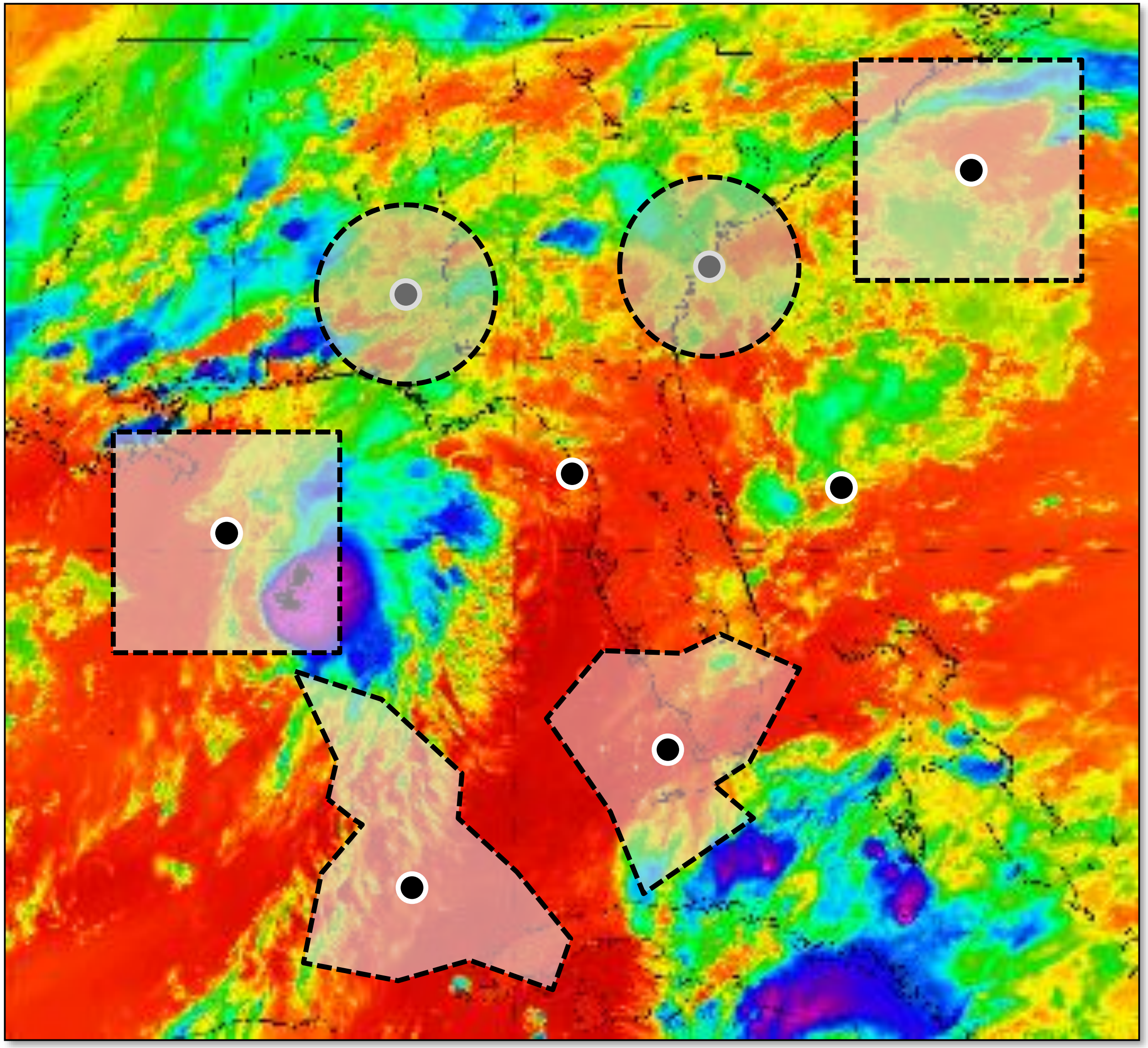


CRU

NOAA

Two possibilities

- Characterize sampling points with the underlying value for a given environmental variable (typically the case for plants)
- Characterize sampling points with statistics of a given environmental variable for the underlying area (typically the case for animals $\rightarrow$ estimation of the movement range)



Environmental representativeness

- We just considered spatial representativeness of environmental variables
- Another issue is the temporal representativeness of environmental variables
- CRU data or Worldclim data are produced on the basis of observations collected during several years (e.g. 30 years for CRU, ~1960-1990)
- Can we compare genetic data which are the result of long evolutionary processes with environmental data characterizing only the very recent period ?
- Paleo-environmental datasets are available

Paleoenvironmental data

ECOGRAPHY
PATTERN AND PROCESS IN ECOLOGY

Explore this journal >

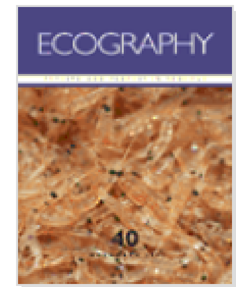
Software note

PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales

Damien A. Fordham, Frédérik Saltré, Sean Haythorne, Tom M. L. Wigley,
Bette L. Otto-Bliesner, Ka Ching Chan, Barry W. Brook

First published: 12 May 2017 Full publication history

DOI: 10.1111/ecog.03031 View/save citation



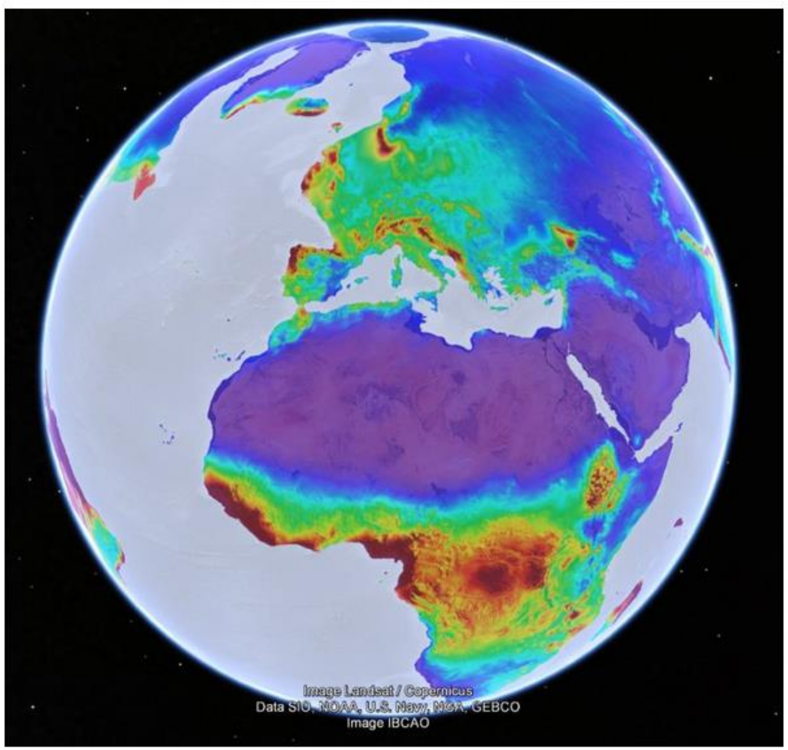
View issue TOC
Volume 40, Issue 11
November 2017
Pages 1348–1358

CHELSA

Climatologies at high resolution for the earth's land surface areas

Last glacial maximum climate

Downscaled global climatological data from the last glacial maximum (21.000 years ago). The CHELSA LGM data is based on a implementation of the CHELSA algorithm on PMIP3 data. For now we only implemented the CCSM4 global circulation model as a test dataset. It is available in via our FILE SERVER.



Annual precipitation patterns during the last glacial maximum. Red areas indicate high precipitation, blue indicates low precipitation.

The data uses a paleoDEM at 30 arc second resolution and an interpolation of

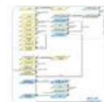
Follow us on Twitter

Tweets von @CHELSA_CLIMATE

Dirk Karger
@CHELSA_CLIMATE

Antwort an @CHELSA_CLIMATE

check out the article in #ScientificData nature.com/articles/sdata... if you are interested in the performance of the timeseries



Climatologies at high resolutio...
High-resolution information on ...
nature.com

26. Okt. 2017

Dirk Karger
@CHELSA_CLIMATE

chelsa-climate.org/timeseries #CHELSA monthly #climate data now online for each month from 1979-2013

Timeseries
The CHELSA timeseries data i...
chelsa-climate.org

25. Okt. 2017

Dirk Karger
@CHELSA_CLIMATE

climatedataguide.ucar.edu/climate-data/c.... #CHELSA at #NCAR #Climate data guide.

25. Okt. 2017

Dirk Karger
@CHELSA_CLIMATE

journals.plos.org/plosone/articl... Using worldwide edaphic data and #CHELSA to model plant

Exercise

- Characterizing sampling points with environmental data