SCIENTIFIC REPORT: EXCHANGE GRANT
Ref Number 3733.

Document describing the following information:
- Purpose of the visit
- Description of the work carried out during the visit
- Description of the main results obtained
- Future collaboration with host institution (if applicable)
- Projected publications/articles resulting or to result from your grant
- Other comments (if any)

ESF activity: LESC; Advances in Farm Animal Genomic Resources; Genomic-Resources
Project title: Molecular dissection of inbreeding depression for bull fertility
Candidate researcher: Maja Ferencakovic, MSc PhD student at the University of Zagreb, Zagreb (Croatia).
Host researcher: Dr. Johann Soelkner, Professor at University of Natural Resources and Applied Life Sciences Vienna, Vienna (Austria)

Purpose of the visit

The main objective of the visit was to analyze inbreeding depression for bull fertility on a molecular level. Inbreeding depression is the reduction of the population mean for a quantitative trait such as size, fertility, vigor, yield, and fitness. Maximini et al. have shown that inbreeding depression does affect male fertility of Simmental bulls using pedigree and sperm quality data. Pedigree data are usually difficult to obtain (impossible for wild animals), they are potentially unreliable, and do not count for inbreeding arising from distant common ancestors. Even if the pedigree is well known and correct the estimates of inbreeding for single individuals can differ from expectation due to the stochastic pattern of recombination. In contrast to pedigree Runs of homozygosity (ROH) were recently proposed as genomic measure quantifying individual autozygosity that counts for stochastic variation of recombination (McQuillan et al., 2008; Nalls et al., 2009). In addition Keller et al. (2011) proposed ROH as optimal way for indentifying ancient inbreeding that we cannot obtain from pedigree data in humans, and Ferencakovic et al. (2011) in cattle. Estimation of precise inbreeding coefficient is important for separating influences of various environmental and physiological factors on quantitative traits, from genetic factors. Genome-wide association studies (GWAS) allow us to scan entire genome for associations between common gene variants (SNPs) and phenotype. Visscher (2008) reported a total of 54 SNPs influencing human height as classical quantitative trait. Furthermore, Pryce et al. (2011) reported orthologous positions of 55 genes associated with height in four human populations located on the bovine genome. Single nucleotide polymorphisms close to eight of these genes were significantly associated with stature in cattle.

To analyze inbreeding depression from molecular perspective our goal was (is) to test next hypotheses:

a) is genomic autozygosity associated to male fertility

b) does the pedigree inbreeding coefficient accurately estimate autozygosity;

c) does ROH inbreeding coefficient gives better insight in inbreeding depression;

d) is inbreeding depression caused by specific chromosomal regions or is it caused by genes from the whole genome.

Description of the work carried out during the visit

Day after my arrival I received data bases with information related to the bull sperm quality (ejaculate volume, sperm concentration, percentage of viable spermatozoa, total number of spermatozoa, motility score) from regular measurements done in Austrian AI (Artificial Insemination) stations. There were approximately 53 000 ejaculates obtained from around 1500 bulls from three breeds (Simmental, Brown Swiss and Tyrol Grey) from the Hohenzell Station from Upper Austria, from Wieselburg, the Station from Lower Austria sampled during six years, from Gleisdorf, the Station from Styria and Birkenberg, the Station from Tyrol. First goal was to clean existing data and use only genotyped animals. This resulted with 1237 animals and 50829 ejaculates

All those animals were genotyped using the Illumina bovine SNP chip with 54 001 SNPs. After quality control we were left with 42262 SNPs that are present in all three populations.

ROH were calculated using PLINK software. From five ROH lengths (>1Mb, >2Mb, >4Mb, >8Mb and >16Mb) inbreeding coefficients (F_{ROH1} , F_{ROH2} , F_{ROH4}, F_{ROH8}, and F_{ROH16}) were calculated as sum of all ROHs of specific length divided by length of genome covered with SNPs. We have also calculated inbreeding for every autosomal chromosome separately using those 5 ROH lengths. For comparison we calculated inbreeding from pedigree for whole pedigree (F_{pedT}) and for five generation pedigree (F_{ped5}) using ENDOG.

Next step was estimation of inbreeding depression. For this purpose we build mixed models with a bull as a random effect and all other effects (age of the bull, semen collector, month and year of collection and number of ejaculates per bull per day), including pedigree and genomic inbreeding coefficients (covariable), as fixed effects. Furthermore, we also used models with chromosomal inbreeding coefficients (1 to 29) as well as models with combination of several chromosomal inbreeding coefficients. Models were then evaluated using Burnham & Anderson method (2002). This last part of analysis we managed to perform only for Simmental bulls from stations Hohenzell and Wieselburg but it is applicable to other breeds and stations and further analyzes are in progress.

Description of the main results obtained

Our preliminary results are giving us some orientation answers on our basic questions and also providing us idea of further work.

a) is genomic autozygosity associated to male fertility

c) does ROH inbreeding coefficient gives better insight in inbreeding depression;

d) is inbreeding depression caused by specific chromosomal regions or is it caused by genes from the whole genome.

Here we are able to conclude that genomic autozygosity is associated to male fertility. Maximini et al. 2010 estimated inbreeding depression using pedigree and sperm quality data on almost same data set as we were. We confirmed their results and found in majority of models presence of inbreeding depression while the strength of models (pedigree versus genomic inbreeding) interchanged depending on a traits or/and population analysed (Table 1-4.) In general, models based on pedigree were comparable to models based on genomic inbreeding (runs of homozygosity) while models based on individual homozygosity (ihom – proportion of homozygous SNPs) were inferior. In contrast, for all traits analyzed, models with various chromosomal inbreeding coefficients outperformed models based on overall genomic inbreeding as well as models based on pedigree. However, while genomic inbreeding coefficients have been shown to be powerful, still much work has left to complete our understanding on architecture of inbreeding depression.

Table 1 Influence of different inbreeding coefficients (covariates) on estimation of inbreeding depression on sperm volume in AI stations Hohenzel and Wieselburg

Volume (ml) Hohenzell					
Covariate	AIC	Δ_null	b	SEE	P
F_{ROH8_chr11}	44930,401	-7,838	-3,906	1,245	0,002
F_{ROH8}	44935,516	-2,724	-9,955	4,613	0,031
F_{ROH2_chr11}	44936,239	-2,000	-2,010	1,012	0,047
F_{ROH2}	44937,174	-1,065	-6,731	3,875	0,082
F_{PED}	44938,127	-0,112	-8,814	6,104	0,149
Null	44938,239	0,000			
F_{PED5}	44939,340	1,101	-6,112	6,491	0,346
ihom	44939,951	1,711	-4,557	8,550	0,594
Volume (ml) Wieselburg					
F_{ROH2}	59298,239	-8,701	-11,592	3,544	0,001
F_{ROH8_chr10}	59301,404	-5,536	-3,423	1,251	0,006
F_{ROH2_chr27}	59301,750	-5,190	-2,631	0,984	0,008
F_{ROH2_chr26}	59301,887	-5,054	-2,251	0,851	0,008
F_{ROH2_chr16}	59302,347	-4,594	-2,514	0,983	0,011
F_{PED}	59302,379	-4,562	-17,491	6,849	0,011
F_{ROH2_chr10}	59302,578	-4,362	-2,806	1,117	0,012
F_{ROH2_chr7}	59303,208	-3,733	-2,654	1,113	0,017
F_{ROH8_chr16}	59304,766	-2,174	-2,511	1,237	0,042
F_{PED5}	59304,869	-2,072	-13,915	6,938	0,045
F_{ROH8}	59305,151	-1,789	-8,869	4,586	0,053
ihom	59305,454	-1,487	-2,376	1,281	0,064
Null	59306,940	0,000			

AIC, Akaike's information criterion ; Δ_null, difference between "Null model" which is model without any inbreeding coefficient and current model (If this value is smaller than -2 then that particular model is better than "Null model". The smaller the value is indicates that this is the model with lowest AIC and the best on); b, regression coefficient; SEE, standard error, P, P value.

Table 2 Influence of different inbreeding coefficients (covariates) on estimation of inbreeding depression on sperm concentration in AI stations Hohenzel and Wieselburg

Concentration (10^9/ml) Hohenzell					
Covariate	AIC	Δ_null	b	SEE	P
F_{ROH8_chr10}	6492,085	-2,358	0,481	0,232	0,038
Null	6494,442	0,000			
F_{ROH8}	6494,793	0,350	1,292	1,014	0,203
ihom	6495,158	0,715	-2,096	1,865	0,261
F_{PED}	6495,190	0,748	-1,478	1,332	0,267
F_{PED5}	6495,685	1,243	-1,225	1,421	0,389
F_{ROH2}	6496,435	1,993	-0,070	0,851	0,934
Concentration (10^9/ml) Wieselburg					
F_{ROH2_chr20}	7113,482	-6,354	-0,506	0,175	0,004
F_{ROH8_chr26}	7114,286	-5,549	0,487	0,178	0,006
F_{ROH2_chr26}	7115,780	-4,055	0,367	0,150	0,014
F_{ROH8_chr19}	7116,091	-3,744	-0,926	0,387	0,017
F_{ROH8_chr22}	7116,270	-3,565	0,435	0,185	0,019
F_{ROH2_chr7}	7117,844	-1,991	-0,392	0,197	0,047
Null	7119,835	0,000			
ihom	7119,915	0,080	0,309	0,224	0,168
F_{PED}	7121,630	1,795	-0,550	1,223	0,653
F_{ROH8}	7121,752	1,916	0,233	0,814	0,775
F_{ROH2}	7121,795	1,960	0,126	0,637	0,844
F_{PED5}	7121,822	1,986	0,143	1,232	0,908

AIC, Akaike's information criterion ; Δ_null, difference between "Null model" which is model without any inbreeding coefficient and current model (If this value is smaller than -2 then that particular model is better than "Null model". The smaller the value is indicates that this is the model with lowest AIC and the best one); b, regression coefficient; SEE, standard error, P, P value.

Table 3. Influence of different inbreeding coefficients (covariates) on estimation of inbreeding depression on sperm viable spermatozoa in AI stations Hohenzel and Wieselburg

Viable spermatozoa (%) Hohenzell					
Covariate	AIC	Δ_null	b	SEE	P
F_{ROH2_chr14}	83790,548	-8,981	-13,555	4,092	0,001
F_{ROH2_chr27}	83795,327	-4,201	-14,057	5,686	0,013
F_{ROH8_chr14}	83796,541	-2,987	-10,620	4,793	0,027
F_{PED}	83797,087	-2,442	-56,042	26,828	0,037
F_{PED5}	83797,352	-2,176	-57,643	28,465	0,043
Null	83799,529	0,000			
F_{ROH2}	83799,592	0,064	-23,669	17,181	0,168
ihom	83800,004	0,475	-46,218	37,830	0,222
F_{ROH8}	83801,492	1,963	-3,934	20,568	0,848
Viable spermatozoa (%) Wieselburg					
F_{ROH8_chr16}	71361,375	-19,174	-12,050	2,627	0,000
F_{ROH2_chr16}	71367,643	-12,906	-8,032	2,091	0,000
Null	71380,549	0,000			
F_{PED}	71380,819	0,270	-18,950	14,544	0,193
F_{PED5}	71380,833	0,284	-19,056	14,683	0,194
F_{ROH2}	71381,142	0,593	-8,901	7,585	0,241
F_{ROH8}	71381,766	1,217	-8,478	9,701	0,382
ihom	71382,515	1,966	-0,498	2,711	0,854

AIC, Akaike's information criterion ; Δ_null, difference between "Null model" which is model without any inbreeding coefficient and current model (If this value is smaller than -2 then that particular model is better than "Null model".The smaller the value is indicates that this is the model with lowest AIC and the best one); b, regression coefficient; SEE, standard error, P, P value.

Table 4. Influence of different inbreeding coefficients (covariates) on estimation of inbreeding depression on sperm motility in AI stations Hohenzel and Wieselburg

Motility Hohenzell					
Covariate	AIC	Δ_null	b	SEE	P
F_{ROH2_chr14}	11024,136	-10,400	-0,408	0,117	0,000
F_{ROH2_chr20}	11027,177	-7,359	-0,347	0,115	0,003
F_{ROH8_chr20}	11030,092	-4,443	-0,415	0,166	0,012
F_{ROH8_chr14}	11032,176	-2,360	-0,279	0,136	0,040
F_{ROH8_chr23}	11032,266	-2,270	0,416	0,205	0,042
F_{ROH2_chr19}	11032,386	-2,150	0,284	0,142	0,046
F_{ROH2_chr27}	11032,459	-2,076	-0,325	0,164	0,047
F_{ROH2_chr23}	11032,484	-2,052	0,263	0,133	0,047
F_{PED5}	11033,683	-0,853	-1,347	0,811	0,097
F_{PED}	11034,100	-0,436	-1,178	0,767	0,124
Null	11034,536	0,000			
F_{ROH2}	11035,199	0,663	-0,567	0,497	0,254
ihom	11035,757	1,222	-0,956	1,101	0,385
F_{ROH8}	11036,463	1,927	-0,161	0,592	0,786
Motility Wieselburg					
F_{ROH8_chr13}	26353,654	-11,680	-2,577	0,695	0,000
F_{ROH2_chr13}	26359,971	-5,362	-1,025	0,380	0,007
F_{ROH8_chr16}	26360,262	-5,071	-0,813	0,308	0,008
F_{ROH2_chr16}	26360,816	-4,517	-0,623	0,246	0,011
F_{PED5}	26365,031	-0,302	-2,603	1,730	0,132
Null	26365,333	0,000			
F_{PED}	26366,458	1,125	-1,588	1,718	0,355
ihom	26366,636	1,303	-0,266	0,322	0,407
F_{ROH2}	26366,895	1,561	-0,585	0,895	0,513
F_{ROH8}	26367,297	1,964	-0,214	1,145	0,852

AIC, Akaike's information criterion ; Δ_null, difference between "Null model" which is model without any inbreeding coefficient and current model (If this value is smaller than -2 then that particular model is better than "Null model". The smaller the value is indicates that this is the model with lowest AIC and the best one); b, regression coefficient; SEE, standard error, P, P value.

b) does the pedigree inbreeding coefficient accurately estimate autozygosity

Here we can conclude that pedigree inbreeding coefficient does not estimate autozygosity accurately. First, this approach fails to capture the influence of relatedness among founders from the base population. Second, F_{PED} is the expected proportion of the genome that is IBD and does not take into account the stochastic nature of recombination. Third, several studies confirm that errors in cattle pedigrees are common due to misinterpretation, misidentification and incorrect recording (e.g., Ron *et al.* 1996). Finally, F_{PED} assumes that the entire genome is selection-neutral and does not account for potential bias resulting from selection (Curik *et al.* 2002). Here we have calculated levels of autozygosity from pedigree data and from genomic data. From genomic data we calculated F_{ROH} and $F_{HOM.}$ F_{HOM} is genomic inbreeding coefficient based on the difference between observed and expected numbers of homozygous genotypes. It is clear from Table 5 that levels of autozygosity are much higher for F_{ROH} and $F_{HOM.}$ F_{PED} and F_{PED5} are similar to values estimated from ROH of 8 or 16Mb which are considered to be from recent inbreeding.

Table 5. Mean, standard deviation and range of calculated inbreeding coefficients in three breeds.

	Brown Swiss			Tyrol Grey			Simmental		
	Mean	Std	Range	Mean	Std	Range	Mean	Std	Range
F_{PED}	0.048	0.020	0.001 - 0.127	0.030	0.024	0.003 - 0.169	0.014	0.013	0.000 - 0.09
F_{PED5}	0.024	0.017	0.000 - 0.106	0.018	0.023	0.000 - 0.159	0.009	0.012	0.000 - 0.085
F_{HOM}	0.138	0.036	0.047 - 0.264	0.076	0.036	0.024 - 0.239	0.066	0.026	0.001 -0.180
F_{ROH1}	**0.151**	**0.032**	0.069 - 0.273	**0.085**	**0.030**	0.041 - 0.234	**0.085**	**0.020**	0.028 -0.183
F_{ROH2}	0.125	0.032	0.046 - 0.252	0.060	0.031	0.017 - 0.213	0.052	0.019	0.009 - 0.150
F_{ROH4}	0.101	0.032	0.029 - 0.227	0.047	0.031	0.008 - 0.203	0.030	0.017	0.002 - 0.124
F_{ROH8}	0.072	0.029	0.012 - 0.194	0.035	0.03	0.000 - 0.183	0.016	0.016	0.000 - 0.102
F_{ROH16}	0.037	0.023	0.000 - 0.154	0.019	0.024	0.000 - 0.140	0.008	0.012	0.000 - 0.062

Future collaboration with host institution (if applicable)

The working groups of Johann Sölkner and Ino Curik have been collaborating on various issues related to inbreeding and inbreeding depression from some time. The availability of high throughput genotypes and male fertility phenotypes allows much deeper insight and tests of some of the hypotheses the two groups had pointed to in previous work. This project represent beginning of future collaboration between these two groups as follow up of our previous work. We have just started to use genomic data as tool for understanding inbreeding and its construction.

Projected publications/articles resulting or to result from your grant

Results from this study will be partly presented at 4th International conference in quantitative genetics 17 - 22 June 2012 Edinburgh, Scotland, UK., and at the 63rd Annual Meeting of the EAAP which will take place in Bratislava from August 27 to August 31, 2012. We are also planning a paper of general interest to be published from our results.