

Prospecting genomic regions associated with milk production traits in Egyptian buffalo

Hamdy Abdel-Shafy, Mohamed A.A. Awad, Hussein El-Regalaty, Salah El-Din El-Assal and Samy Abou-Bakr

Supplementary table 1- Editing steps for daily milk records in the buffalo cows

Editing options	Deleted		Remaining	
	Records	Animals	Records	Animals
Initial numbers	-----	-----	177,995	1,992
Known calving dates	178	15	177,817	1,977
Lactation number between 1 and 13	1,739	3	176,078	1,974
DIM between 5 and 290	10,384	18	165,694	1,956
At least 3 months of lactation	4,215	286	161,479	1,670

DIM: days in milk.

Supplementary table 2- Editing process for SNPs used for GWAS in the Egyptian buffalo

Quality control method	Deleted SNPs	Remaining SNPs
Original SNPs	-----	123,040
Duplicate SNP positions	27,913	95,127
Un-positioned SNPs	17,845	77,282
SNPs with low call rate <85%	7,045	70,237
SNPs deviated from HWE	823	69,414
SNPs with MAF<0.01	5,245	64,169
LD pruning ($r^2>0.5$)	19,184	44,985

MAF: minor allele frequency, HWE: Hardy-Weinberg equilibrium, LD: linkage disequilibrium

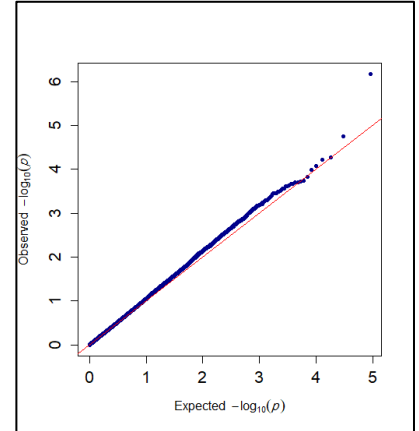
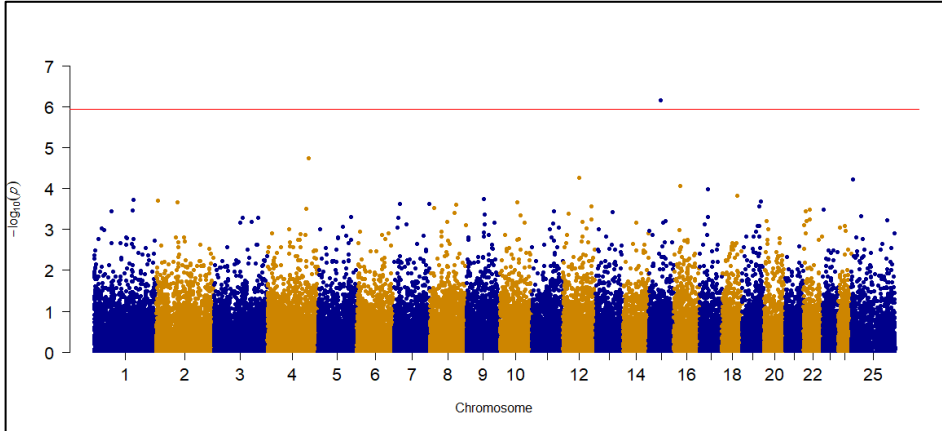
Supplementary table 3- Suggestive SNPs and candidate genes associated with milk production traits in Egyptian buffalo

Traits	SNP ID	Chr.	Positions [bp]	MA	MAF	β	P-value	Nearest gene	Distance	Other genes	Distance
DMY	AX-85071608	4	135416124	C	0.01	-1.411	1.75E-05	ANK3	Intron		
FP	AX-85125496	15	48437751	T	0.03	-1.246	1.97E-06	NCOA2	Intron		
FP	AX-85107786	23	50923849	G	0.01	0.930	4.00E-06	ADGRA1	29570	CFAP46	72007
FP	AX-85043945	25	118884025	G	0.01	1.102	6.85E-06	LOC102391418	Intron		
FY	AX-85124959	7	11467929	C	0.02	0.149	4.83E-06	HS3ST1	362436	CLNK	410664
FY	AX-85055593	15	38842902	T	0.02	0.154	6.69E-06	TPD52	114860	ZBTB10	189227
FY	AX-85111822	15	2960480	G	0.01	0.208	6.93E-06	RALYL	771376	SNX16	1489837
FY	AX-85118863	5	60745014	G	0.02	0.208	6.93E-06	ESRRG	Intron		
FY	AX-85120516	9	106373547	T	0.05	0.075	7.88E-06	MAU2	Intron		
FY	AX-85065545	1	174771819	G	0.16	-0.046	9.53E-06	FOXL2	-83881	PIK3CB	22585
PP	AX-85043902	4	73662085	G	0.02	0.263	1.25E-06	GRIP1	547020	CAND1	94439
PP	AX-85101450	2	15690170	T	0.04	-0.127	1.42E-06	CDKAL1	Intron		
PP	AX-85075927	8	30164345	A	0.02	-0.348	1.62E-06	DNAH11	Intron		
PP	AX-85125519	17	61542442	T	0.08	-0.106	2.03E-06	TTC29	Intron		
PY	AX-85120516	9	106373547	T	0.05	0.033	2.56E-06	MAU2	Intron		
PY	AX-85126095	14	70369288	A	0.02	0.037	5.00E-06	GATA3	-404324	CELF2	2063859
PY	AX-85043827	2	57503555	G	0.01	0.052	5.38E-06	IWS1	Intron		
PY	AX-85064305	24	41783140	A	0.02	0.056	5.55E-06	LMF1	Intron		
PY	AX-85043902	4	73662085	G	0.02	0.066	7.52E-06	CAND1	94439	GRIP1	547020
PY	AX-85052771	14	82284533	G	0.43	0.016	8.34E-06	PLCB1	Intron		
PY	AX-85077812	3	158539332	A	0.23	-0.018	9.22E-06	TMEM38B	-375025	ZNF462	682788
PY	AX-85103359	19	41292383	A	0.07	0.031	9.61E-06	PDZD2	Intron		
PY	AX-85124959	7	11467929	C	0.02	0.060	1.11E-05	HS3ST1	-362436	CLNK	410664
PY	AX-85090818	18	40134051	T	0.11	0.023	1.34E-05	CCNE1	46481	C18H19orf12	46613

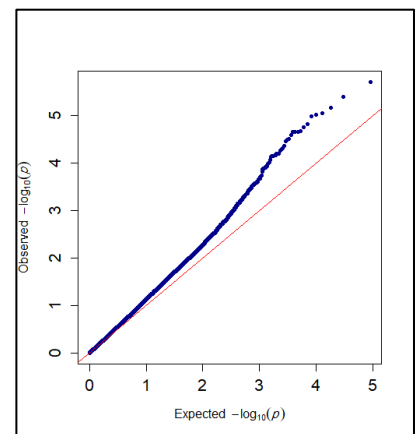
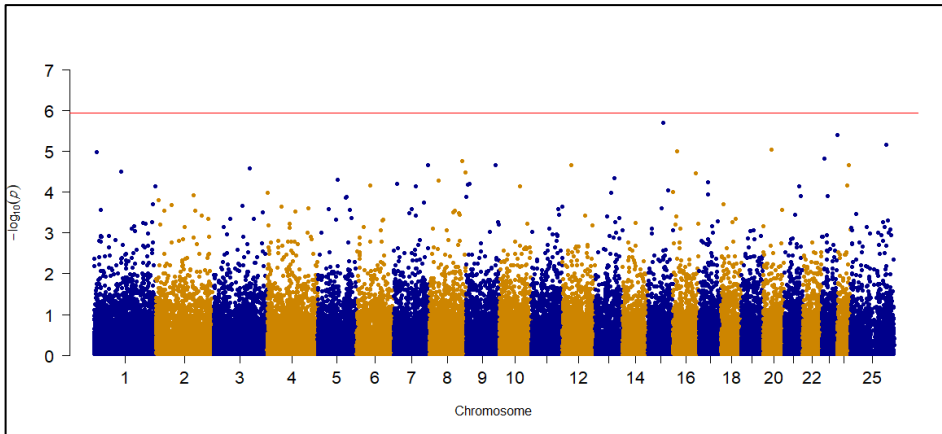
Chr: chromosome, MA: minor allele, MAF: minor allele frequency, β : change per minor allele, DMY: daily milk yield, DFY: daily fat yield, DPY: daily protein yield, DLY: daily lactose yield. Suggestive SNPs are close to significant threshold line based of Bonferroni correction. In the distance: "+" for upstream and "-" for downstream. Positions are given according to the latest reference assembly of buffalo genome (UOA_WB_1: GCA_003121395.1).

Supplementary figure 1 – Manhattan and Q-Q plots of genome-wide association for a) daily milk yield, b) fat%, c) fat yield, d) protein%, and e) protein yield in Egyptian buffalo.

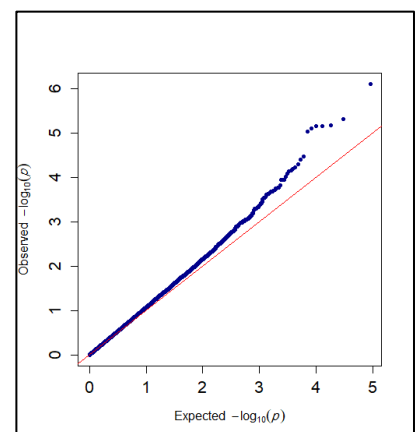
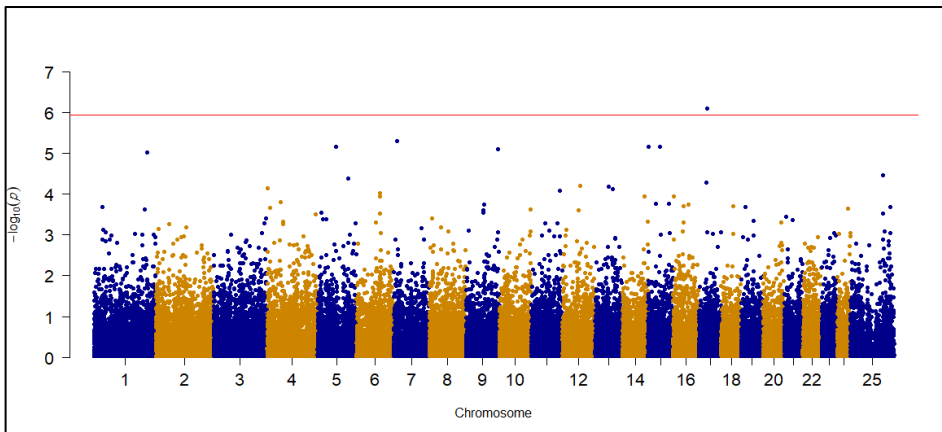
a)



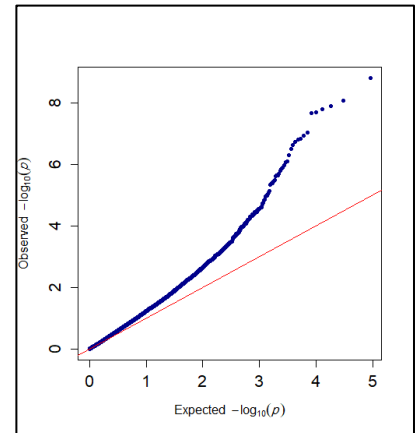
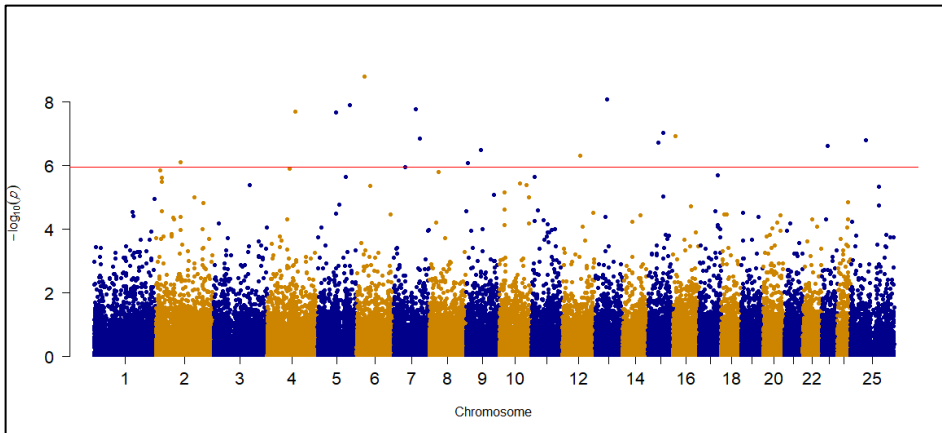
b)



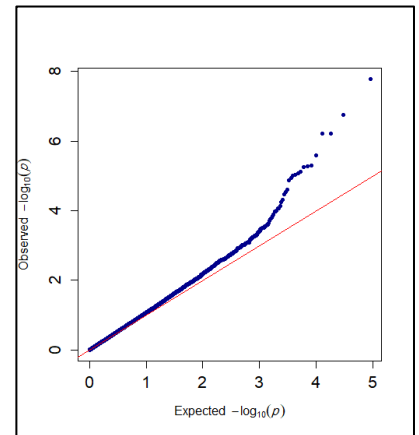
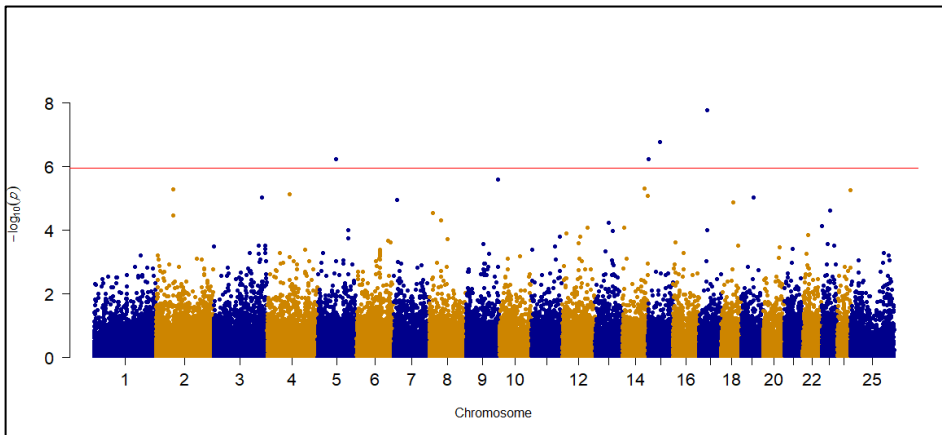
c)



d)



e)



The horizontal line indicates the whole-genome significance threshold after Bonferroni correction at $\alpha = 0.05$ ($P \leq 1.11 \times 10^{-6}$).