Archiv Tierzucht 56 (2013) 69, 691-699

Short communication

The effects of artificial selection on genetic variation of some immune genes in *Gallus gallus*

Tangjie Zhang, Nannan Zhao and Qian Liu

College of Veterinary Medicine, Yangzhou University, Yangzhou, China

Abstract

To research effects of the artificial selection of *Gallus gallus* on *G. domesticus*' nucleotide diversity of immune genes, sequence polymorphisms of *G. domesticus* (23 genes), *G. gallus* (23 genes), *G. lafayetti* (17 genes), and *G. sonneratii* (17 genes) were obtained from GenBank. The data set included 819 polymorphisms. Immune gene polymorphism and selection efficiency in the data from those four species of *Gallus* were calculated. By calculating the q_w (Watterson's estimator) of each site, an average q_w for each species and the minimum number of re-combinations in each species and by estimating the selection efficiency for *G. domesticus* and *G. gallus*, neither significant nucleotide diversity nor genetic-diversity-q_w-difference was found between *G. domesticus* and *G. gallus*. The results indicated that the patterns of genetic diversity in *G. domesticus* were strongly influenced by recombination and, because Tajima's D has a negative value, recombination was the main mechanism responsible for the immune gene evolution of *G. gallus*.

Keywords: domestication, artificial selection, Gallus gallus, immune gene

Introduction

Domestic animals have often been artificially selected for certain traits over several thousand years. Poultry domestication is the genetic modification of a wild species to create a new form of a bird to meet human needs. Improvement after domestication has also resulted in striking changes in yield, immune system, biochemical composition and other traits.

The domestic chicken is descended primarily from the Red Junglefowl (*G. gallus*) in Southeast Asia nearly 10 000 years ago (Crawford 1990). But at least one other species

Archiv Tierzucht 56 (2013) 69, 691-699
doi: 10.7482/0003-9438-56-069

Received: 07 May 2012
Accepted: 11 June 2013
Online: 14 June 2013

Corresponding author:
Tangjie Zhang; email: slx@yzu.edu.cn
College of Veterinary Medicine, Yangzhou University, Yangzhou 225009, China

must have contributed, specifically the grey jungle fowl (Wong *et al*. 2004), to the domestic chicken. The size, shape and production of the modern domesticated chicken have been sculpted by artificial selection for at least 2 000 years, likely contain many important lessons about the genetic architecture of phenotypic variation and the mechanistic basis of selection. Indeed, chicken and other domesticated species played an important role in Darwin's »On the Origin of the Species«, as they provided vivid examples of descent with modification.

Most domesticated animals have experienced »a domestication bottlenec« that reduced genetic diversity relative to their wild ancestor (Buckler 2001). This bottleneck affects all genes in the genome and modifies the distribution of genetic variation among loci. Selection is similar to a more severe bottleneck (Galtier 2000) that removes most of the genetic variation from a target locus. Chicken (*G. domesticus*) showed a high density of SNP and a high recombination rate, which made it possible to perform high-throughput genotyping to evaluate the existing genetic diversity in chicken at the genome level compared to other species. However, relatively little progress has been made on systematically identifying which immune gene sites of *G. domesticus* genome were influenced by selective breeding during the natural history of chicken.

Here, genetic variations of 23 gene fragments in a sample of *Gallus* Genus 4 species, *G. domesticus*, *G. gallus*, *G. lafayetii*, *G. sonneratii* on the basis of gene sequence polymorphism were reported and the effects of artificial selection on some immune genes were analysed in *G. gallus*. The multi-locus analysis is a powerful way to detect adaptation at the population level, so that comparative researches of diversity and recombination in the *Gallus* genus would help us to comprehensively understand the immune genetic structure and selection in domestication.

Material and methods

Gallus families sequence polymorphism data set

A total of 819 data sets of *Gallus* genus gene polymorphisms (1 159 to 9 398 base pairs) was obtained from Popset of GenBank and Daniel G. Bradley (2010) including sequence polymorphisms of *G. domesticus* 23 genes, *G. gallus* 23 genes, *G. lafayetii* 17 genes and *G. sonneratii* 17 genes. Each group was aligned by eye using CLUSTALW (Thompson *et al.* 1994). Alignments of all groups are available on request.

Polymorphism sequence data analyses

The average variability Pi and minimum number of the recombination parameter were calculated by using DnaSP v. 5 (Librado & Rozas 2009). Insertions/deletions (indels) were excluded from all estimates. To investigate the evidence of the non-neutral evolution, the D test of Tajima was applied (Tajima 1989).

The diversity was measured by Watterson's estimator of the population mutation parameters (q_w), which was calculated separately from the non-synonymous and silent sites for the con-specific gene fragments. The parameter represents the per-site diversity.

$$q_w = \frac{P}{L\sum_{i=1}^{n-1}\frac{1}{i}} \quad (1)$$

P is the number of synonymous polymorphisms, *L* is the number of synonymous sites and n is the number of the sequence sampled.

We calculated the efficiency of selection for four species. Measurement of selection efficiency is as following:

$$\frac{q_n}{q_{s+i}} = \frac{\sum_n P_n / \sum_n L_n}{(\sum_s P_s + \sum_i P_i + 1)/(\sum_s L_s + \sum_i L_i)} \quad (2)$$

where P_n, P_s and P_i are the numbers of non-synonymous, synonymous and intron polymorphisms. L_n, L_s and L_i are the numbers of non-synonymous, synonymous and intron sites for each gene in each species. q_n is for non-synonymous sites. q_{s+i} is for synonymous and intron correspondingly. *q* were arc-sine transformed. Recombination parameter was log x+1 transformed. After the calculation of q_s for synonymous sites and q_i for intron sites, the weighted average of q_s or/and q_i from different genes for same species was made.

Results

Whole sequence segment variations in G. domesticus and G. gallus

More than 84 kb of the DNA sequence was obtained across 23 immune genes. Sequence data for the *CR1* and *OTC* gene were obtained from only *G. domesticus* and *G. gallus*. Summary statistics of the number of segregating sites, nucleotide diversity Pi, q_w, Tajima's D test and the minimum number of recombination events were shown in Table 1. Single variable regressions of immune gene nucleotide diversity showed that no significant difference was found between *G. domesticus* and *G. gallus* ($F_{2,23}$=1.39, *P*=0.5845>0.05). Neither did the diversity-q_w of the fragments ($F_{2,22}$=0.074, *P*=0.7866>0.05).

The population recombination parameter, ρ, is the other key parameter in simple population genetic models. However, the estimation of ρ requires considerably larger segments of contiguous DNA to be sequenced (Hudson 2001). The relatively short sequences obtained in this study are not sufficient to provide reliable locus-specific estimates of ρ. Instead, the estimation of the minimum number of the recombination parameter was obtained from the two sample species. Single variable regressions revealed that the recombination parameter of *G. domesticus* was significantly higher than that of *G. gallus* ($F_{2,23}$=6.160, *P*=0.0169<0.05) after recombination parameter was log x+1 transformed. By using DnaSP, recombination parameter was estimated, which is inversely proportional to LD (linkage disequilibrium). The average of estimates of recombination parameter in *G. domesticus* is 164 % of that in *G. gallus*, while the average of estimates of q_w in *G. domesticus* is 95 % of that in *G. gallus* (Figure 1). Thus, the recombination parameter in *G. domesticus* has been reduced more drastically than the population mutation parameter q_w, contrary to what has been expected under a population

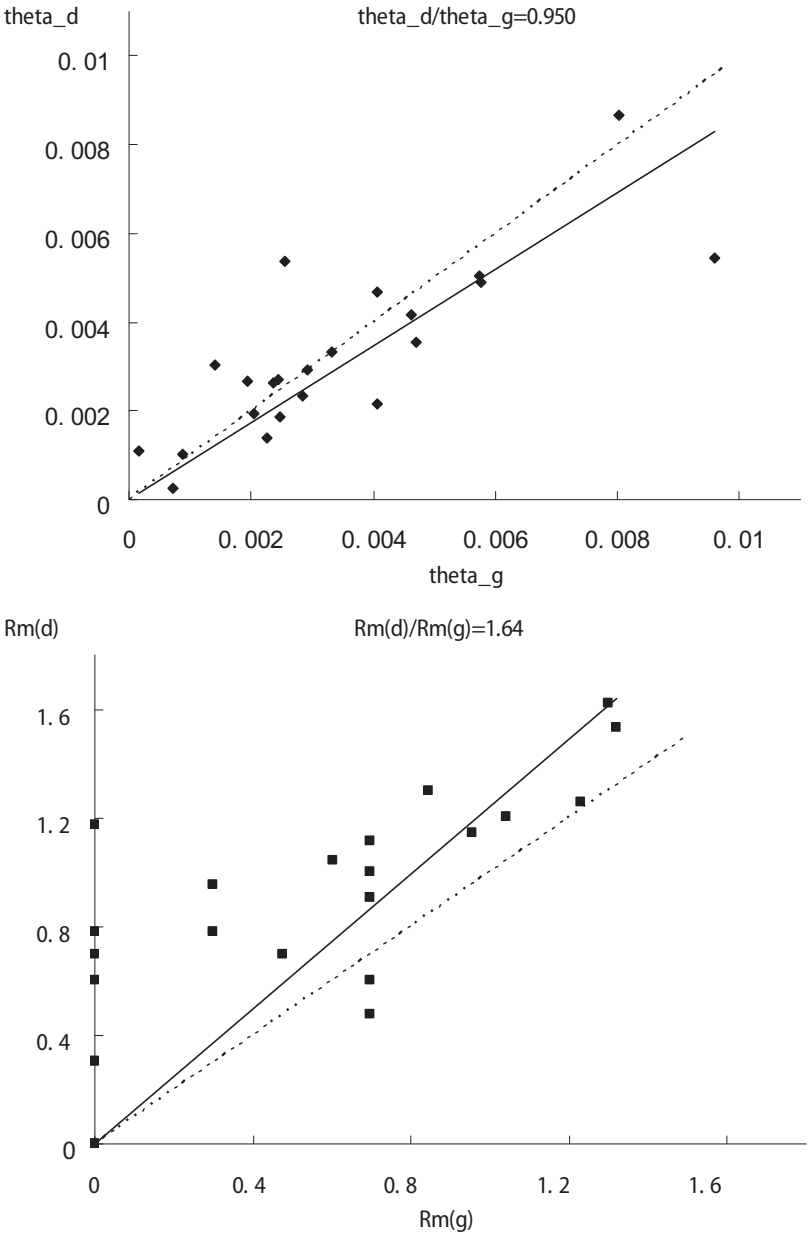
Table 1
Pi, Theta, TajimaD and recombination parameter of 23 gene sequence for 4 species

Gene	L	Ns *G. d.*	Ns *G. g.*	N *G. d.*	N *G. g.*	Pi *G. d.*	Pi *G. g.*	q_w *G. d.*	q_w *G. g.*	D *G. d.*	D *G. g.*	Rm *G. d.*	Rm *G. g.*
KK34	4 381	17	10	26	8	0.001126	0.000858	0.001019	0.00088	0.3702	-0.137	5	1
TLR15	3 105	21	20	21	8	0.001886	0.002059	0.001855	0.00248	0.0625	-0.8914	9	4
TLR1LA	3 559	41	13	26	8	0.00168	0.001726	0.003019	0.00141	-1.679	1.1403	4	2
TLR1LB1	6 047	80	52	30	8	0.0032	0.003874	0.003339	0.00332	-0.1587	0.9058	33	20
TLR1LB2	9 398	41	4	30	8	0.000692	0.000179	0.001101	0.00016	-1.3716	0.3842	5	0
TLR2B	4 186	4	7	30	8	0.000328	0.000687	0.000266	0.00071	0.6009	-0.1632	1	0
TLR3	6 628	123	70	30	8	0.004168	0.004117	0.004684	0.00407	-0.4229	0.0579	41	19
TLR4	6 237	58	46	30	8	0.002611	0.002938	0.002347	0.00284	0.4222	0.1757	17	16
TLR5	3 150	59	8	19	2	0.002696	0.00254	0.005359	0.00254	-2.035**	na	4	0
TLR7	6 746	71	34	30	8	0.001989	0.001853	0.002657	0.00194	-0.9503	-0.2489	13	8
GMCSF	2 730	15	16	30	8	0.001929	0.002272	0.001392	0.00227	1.2985	0.0065	3	4
IL13	2 794	38	34	27	8	0.004942	0.004883	0.003529	0.00469	1.4994	0.2151	14	0
IL12A	2 005	33	24	30	8	0.004446	0.005825	0.004155	0.00462	0.2556	1.3747	12	4
IL9	2 954	57	44	26	8	0.003885	0.005356	0.005057	0.00575	-0.8904	-0.3629	15	10
IL8	3 114	32	19	28	8	0.002309	0.002523	0.002641	0.00235	-0.4617	0.375	7	4
IL5	4 686	34	20	30	8	0.002423	0.002506	0.002712	0.00244	-0.389	0.1452	8	1
IL4	1 998	17	21	30	8	0.003122	0.00395	0.002148	0.00405	1.5513	-0.133	10	3
IL3	5 505	42	29	30	8	0.002675	0.001787	0.001929	0.00204	1.4267	-0.6452	19	6
CR1_1	458	2	2	3	3	0.002911	0.002911	0.002911	0.00291	na	na	0	0
CR1_2	499	9	6	5	3	0.008818	0.008016	0.008657	0.00802	0.1316	na	3	0
CR1_3	600	6	12	4	5	0.005	0.008333	0.005455	0.0096	-0.8086	-0.9543	0	0
CR1_4	1 023	7	5	6	2	0.003739	0.00353	0.003244	na	0.8878	na	1	0
OTC	2 210	24	34	6	6	0.004988	0.006345	0.004884	0.00577	0.1331	-0.4086	2	4
Mean						0.003	0.003	0.00323	0.0034	-0.023	0.036		

G. d.: Gallus domestiucs, G. g.: Gallus gallus, L: length of gene, Ns: number of segregating sites, D: Tajima's D, Rm: recombination parameter, na: not available

Table 2
Different sites' q_w, an average q_w and efficiency of selection for 4 species

	G. domesticus			*G. gallus*			*G. lafayetii*			*G. sonneratii*		
Gene	q_s	q_n	q_{s+i}	q_s	q_n	q_{s+i}	q_s	q_n	q_{s+i}	q_s	q_n	q_{s+i}
KK34	0.01184	0.00331	0.00079	0.014	0.00097	0.00087	0	0	0.00027	0	0.00137	0.00027
TLR15	0.00459	0.00109	0.00325	0.0052	0.00124	0.00355	0.003	0.00081	0.00199	0.0091	0.00217	0.00897
TLR1LA	0.00234	0.00263	0.00226	0.0014	0.00102	0.00186				0	0.00058	0.00066
TLR1LB1	0.00171	0.00183	0.00384	0.0026	0.00153	0.00392	0	0	0.00066	0	0.00066	0.00221
TLR1LB2	0.00296	0.001	0.00109	0	0.00084	0	0	0.00059	0.00014	0.0019	0.00108	0.00013
TLR2B	0.00051	0.00032	0.00033	0.0008	0.00023	0.00111	0.001	0	0.00078	0.002	0.00059	0.00239
TLR3	0.00629	0.00291	0.00441	0.0062	0.00374	0.00416	0.006	0.0025	0.00433	0.0059	0.00166	0.00449
TLR4	0.00511	0.00117	0.00288	0.0072	0.00119	0.00359	0.01	0.00169	0.00266	0.0083	0.00113	0.00292
TLR5	0.00462	0.00121	0.0055	0.0051	0.002	0.00348	0.016	0.00464	0.01374			
TLR7	0	0	0.00418	0	0.00016	0.00297	0.001	0.00041	0.0028	0.0008	0.00022	0.00572
GMCSF	0.00245	0.00077	0.00148	0.0075	0	0.00258	0.005	0	0.00251	0	0	0.0016
IL13	0.00738	0.00083	0.00387	0.011	0.00124	0.00513	0	0	0.00439	0.0052	0	0.00352
IL12A	0.00225	0.00304	0.00437	0.0034	0	0.00553	0	0	0.00065	0.0097	0.00329	0.00815
IL9	0.00553	0.00163	0.00548	0.0041	0	0.00645	0	0	0.00608	0	0	0.00995
IL8	0	0	0.00286	0	0	0.00255	0	0	0.00095	0.0071	0	0.00455
IL5	0.00543	0.0037	0.00267	0	0	0.00255	0	0	0.00216	0	0	0.00162
IL4	0.00265	0.00084	0.00233	0.0039	0.0037	0.00412	0	0	0.00259	0	0.00175	0.00194
IL3	0	0	0.00205	0	0	0.00219	0	0	0.00074	0	0	0.00074
CR1_1			0.00291			0.00291						
CR1_2			0.00866			0.00802						
CR1_3			0.00545			0.0096						
CR1_4	0.00367	0.0031	0.00367	0.0056	0.00283	0.0056						
OTC		0.00488			0.00577							
Weighted		0.00173	0.00287		0.00171	0.00293		0.00117	0.00216		0.00094	0.00290
q_n/q_w		0.3760			0.3685			0.3514			0.2448	
q_n/q_{s+i}		0.60192			0.58515			0.54188			0.32559	



The first row illustrates the relationship between mean values of q_w in *G. domesticus* (y-axis) versus *G. gallus* (x-axis). Dashed diagonal lines have a slope of 1.0, representing equal diversity between taxa; solid lines lines are regression lines. Each square represents a single gene.
The second row plots the relationship between estimates of minimum number of recombination in *G. domesticus* (y-axis) versus *G. gallus* (x-axis).

Figure 1
Patterns of diversity in *G. domesticus* and *G. gallus* at 23 gene fragments

bottleneck (Wall *et al.* 2002). The results suggest that patterns of immune gene diversity in *G. domesticus* are strongly influenced by recombination.

Tajima's D test is to distinguish between a DNA sequence evolving neutrally and DNA evolving under a non-random process, including selection, demographic expansion or contraction. In order to perform the test, homologous DNA for at least three individuals was required, so that five sequence polymorphism groups were not available (Table 1). In terms of 23 sampled genes, *TLR5* in *G. domesticus* was found to be statistically significant for Tajima's D ($P<0.05$). In principle, this could potentially indicate a deviation from neutrality, possibly due to strong selection. An average negative Tajima's D in *G. domesticus* (−0.023), in contrast to *G. gallus* (0.036), signifies slightly more low frequency polymorphisms, indicating a population size expansion and/or selection in *G. domesticus*.

Variation of different sites of gene sequence and efficiency of selection in Gallus genus

Across the 23 genes, 84 013 base pairs were aligned, including 831 mutation sites. The comparison of different sites' $q_{w'}$ a weighted average of q_w and selection efficiency for four species was shown in Table 2. The weighted average of q_n for non-synonymous sites was arranged as follows: *G. domesticus* (0.00173), *G. gallus* (0.00171), *G. lafayetii* (0.00117) and *G. sonneratii* (0.00094). In the top-to-bottom order, while the difference of the weighted average q_{s+i} for four species was not found, the arrangement was as follows: *G. domesticus* (0.00287), *G. gallus* (0.00293), *G. lafayetii* (0.00216) and *G. sonneratii* (0.00290), respectively. q_n/q_w represented the percent of mutations of the non-synonymous in total gene segment mutations. *G. domesticus* showed the highest value of q_n/q_w (0.60192) for four species analysed and suggested that *G. domesticus'* immune genes may undergo more strong pressure of selection.

Discussion

In this paper, sequence segment variations, diverse sites weighted mutations (q_w), recombination parameter and efficiency of selection within *Gallus* genus were determined. The results demonstrated the artificial selection of *G. domesticus* and indistinguishable immune genetic diversity with three other species.

A diversity index (Pi) analysis indicated that the overall nucleotide variability of all 23 immune genes for *G. domesticus* and *G. gallus* were approximately 0.003 and showed no difference, as mentioned above. The population recombination rate ρ is a fundamental parameter for evolutionary biology. Not only recombination is a key force shaping the architecture of genomes, but also distribution across genomic regions is essential for association studies of traits. However, the estimation of the population recombination rate is not an easy task. Adequate and reliable locus-specific estimates of ρ could not be provided with relatively short sequences, as we failed in estimation by LDhat v2.0 (McVean 2004), a package for the analysis of recombination rates from population genetic data. We turned to calculate the minimum number of the recombination parameter. The result revealed that the recombination parameter in *G. domesticus* was higher than that in *G. gallus* in most of the sampled loci (21/23).

Of 84 013 base pairs, 831 mutation sites were found. This is somewhat more than the extensive sequence diversity present in domestic chicken (~5 single nucleotide polymorphisms per kilobase in pairwise comparisons) (Wong *et al.* 2004), mainly as a result of the sample size.

Regions of intergenic, noncoding DNA where levels of variation are expected to be higher (Zwick *et al.* 2000) may provide a different picture of diversity. Our estimates of the weighted average of q_w for both silent mutation sites and non-synonymous sites indicated that q_w for silent mutation sites in total species were higher than that for non-synonymous sites, as we expected. *G. lafayetii* has the lowest synonymous mutation q_w=0.00216 among four species, which is obviously correlated to its effective population size.

The mutation parameter of non-synonymous sites-q_n-and percent of non-synonymous mutations in total segment mutations-q_n/q_w-for *G. domesticus* and *G. gallus* were higher than that for *G. lafayetii* and *G. sonneratii*, which indicated that the immune genes of *G. domesticus* and *G. gallus* could undergo a stronger directional selection pressure. This was inconsistent with previous researches that selection for body weight in chicken has depressed immune performance (Miller *et al.* 1992) and antibody production (Cheema *et al.* 2003). Now that almost equivalent q_w for *G. domesticus* and *G. gallus*, there should be some factors which could decrease nuclear diversity of *G. domesticus* since this species obviously have high effective population size, Ne. Of these factors, high recombination event in *G. domesticus* was inferred to be essential.

The highest efficiency of selection (0.60192) was found in *G. domesticus* (Table 2). This selection was mainly described as artificial selection for the needs of human being here, which could be confirmed by negative noticeable Tajima's D.

A population bottleneck was not found by analysis of mutation and recombination parameter. Nevertheless, *G. domesticus* did experience severe population bottleneck (Mason 1984) although this bottleneck effect did not result in a substantial loss of genetic diversity. Abroad crossing between breeds and higher recombination could be fundamental explanations for undifferentiated diversity between *G. domesticus* and *G. gallus*.

Of particular interest will be to define the number of loci responsible for shaping the diversity of form and function, the types of genes and genetic variation therein that have responded to artificial selection. Although our results did not provide definitive answers to these issues, they did afford some insight into the mechanistic basis of artificial selection. Despite the insights gleaned from our data, one limitation of this study was that it did not provide information about more gene sequence polymorphisms. The difference in part reflected differences in sampling.

Acknowledgements

This research was supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions and Jiangsu Co-innovation Center for Prevention and Control of important Animal Infectious Diseases and Zoonoses, Yangzhou, 225009, China.

References

Buckler IV ES, Thornsberry JM, Kresovich S (2001) Molecular Diversity, Structure and Domestication of Grasses. Genet Res 77, 213-218

Cheema MA, Qureshi MA, Havenstein GB (2003) A comparison of the immune response of a 2001 commercial broiler with a 1957 randombred broiler strain when fed representative 1957 and 2001 broiler diets. Poult Sci 82, 1519-1529

Crawford RD (1990) Origin and history of poultry species. In: Crawford RD (ed.) Poultry Breeding and Genetics. Elsevier, Amsterdam, the Netherlands, 1-41

Darwin CR (1872) The Origin of Species. 6th ed. Murray, London, UK

Galtier N, Depaulis F, Barton NH (2000) Detecting Bottlenecks and Selective Sweeps From DNA Sequence Polymorphism. Genetics 155, 981-987

Hillier LW, *et al*. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695-716

Hudson RR (2001) Two-Locus Sampling Distributions and Their Application. Genetics 159, 1805-1817

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452

Mason IL (1984) Evolution of domesticated animals. Longman, London *et al*., UK

McVean G. (2004) LDhat: A package for the population genetic analysis of recombination. Oxford, UK

Miller LL, Siegel PB, Dunnington EA (1992) Inheritance of Antibody Response to Sheep Erythrocytes in Lines of Chickens Divergently Selected for Fifty-Six-Day Body Weight and Their Crosses. Poult Sci 71, 47-52

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585-595

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22 , 4673-4680

Wall JD, Andolfatto P, Przeworski M (2002) Testing Models of Selection and Demography in *Drosophila simulans*. Genetics 162, 203-216

Wong GKS *et al*. (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. Nature 432, 717-722

Zwick ME, Cutler DJ, Chakravarti A (2000) Patterns of Genetic Variation in Mendelian and Complex Traits. Annu Rev Genomics Hum Genet 1, 387-407