

Research article

Open Access

## Estimates of linkage disequilibrium and effective population size in rainbow trout

Caird E Rexroad III\* and Roger L Vallejo

Address: USDA/ARS National Center for Cool and Cold Water Aquaculture, Leetown, West Virginia, 25430, USA

Email: Caird E Rexroad\* - [caird.rexroadiii@ars.usda.gov](mailto:caird.rexroadiii@ars.usda.gov); Roger L Vallejo - [roger.vallejo@ars.usda.gov](mailto:roger.vallejo@ars.usda.gov)

\* Corresponding author

Published: 14 December 2009

Received: 8 September 2009

BMC Genetics 2009, 10:83 doi:10.1186/1471-2156-10-83

Accepted: 14 December 2009

This article is available from: <http://www.biomedcentral.com/1471-2156/10/83>

© 2009 Rexroad and Vallejo; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The use of molecular genetic technologies for broodstock management and selective breeding of aquaculture species is becoming increasingly more common with the continued development of genome tools and reagents. Several laboratories have produced genetic maps for rainbow trout to aid in the identification of loci affecting phenotypes of interest. These maps have resulted in the identification of many quantitative/qualitative trait loci affecting phenotypic variation in traits associated with albinism, disease resistance, temperature tolerance, sex determination, embryonic development rate, spawning date, condition factor and growth. Unfortunately, the elucidation of the precise allelic variation and/or genes underlying phenotypic diversity has yet to be achieved in this species having low marker densities and lacking a whole genome reference sequence. Experimental designs which integrate segregation analyses with linkage disequilibrium (LD) approaches facilitate the discovery of genes affecting important traits. To date the extent of LD has been characterized for humans and several agriculturally important livestock species but not for rainbow trout.

**Results:** We observed that the level of LD between syntenic loci decayed rapidly at distances greater than 2 cM which is similar to observations of LD in other agriculturally important species including cattle, sheep, pigs and chickens. However, in some cases significant LD was also observed up to 50 cM. Our estimate of effective population size based on genome wide estimates of LD for the NCCCWA broodstock population was 145, indicating that this population will respond well to high selection intensity. However, the range of effective population size based on individual chromosomes was 75.51 - 203.35, possibly indicating that suites of genes on each chromosome are disproportionately under selection pressures.

**Conclusions:** Our results indicate that large numbers of markers, more than are currently available for this species, will be required to enable the use of genome-wide integrated mapping approaches aimed at identifying genes of interest in rainbow trout.

### Background

The use of molecular genetic technologies for broodstock management and selective breeding of aquaculture species is becoming increasingly more common with the con-

tinued development of genome tools and reagents for species of interest [1]. Rainbow trout are the most widely produced salmonid in the US, attracting significant interest due to their economic impacts as an aquaculture spe-

cies and on sport fisheries, and as a model research organism for studies related to carcinogenesis, toxicology, comparative immunology, disease ecology, physiology and nutrition [2]. To this end several international laboratories have produced genetic maps for this species to aid in the identification of loci affecting phenotypes of interest. These maps primarily include amplified fragment length polymorphisms (AFLPs) and microsatellites [3-9] and have resulted in the identification of many quantitative/qualitative trait loci (QTL) affecting phenotypic variation in traits associated with albinism, disease resistance, temperature tolerance, sex determination, embryonic development rate, spawning date, condition factor and growth [10-19]. In spite of these efforts, the elucidation of the precise allelic variations and/or genes underlying phenotypic diversity has yet to be achieved in this species having low marker densities and lacking a whole genome reference sequence.

Experimental designs which integrate complex segregation analyses with linkage disequilibrium (LD) approaches facilitate the discovery of genes affecting important traits [20-22]. To this end, the extent of LD has been characterized for humans and several agriculturally important livestock species including cattle, sheep, chickens, and pigs. Farnir *et al.* [20] genotyped 284 autosomal microsatellite markers on 581 dutch black-and-white dairy cattle to construct a whole genome LD map (excluding the sex chromosomes) spanning 2702 cM. Estimations of LD between syntenic loci using Lewontin's normalized  $D'$  [23] revealed large blocks of LD spanning tens of centiMorgans including values of 50% for markers <5 cM and decaying to 16% for distances of 50 cM. The value of  $D'$  between non-syntenic loci was estimated to be 12%. Vallejo *et al.* [24] selected distantly related animals to quantify the level of genetic diversity in United States Holstein cattle. While only 23 Holstein bulls were genotyped with 54 microsatellite loci that spanned most of the autosomal genome, extensive LD was detected in the United States Holstein population in agreement with the findings of Farnir *et al.* [20]. In 2002, McRae *et al.* [25] observed similar results by genotyping 90 microsatellites from 10 chromosomes on 276 progeny from Coopworth sheep, estimating  $D'$  values of 34.3% for marker distances <60 cM, 12.4% for syntenic markers >60 cM, and 12.4% for non-syntenic loci. In 2005 Heifetz *et al.* [26] reported that significant LD was only observed for loci < 5 cM apart in a commercial layer chicken population. In 2006 Harmegnies *et al.* [27] characterized LD in two commercial pig populations of 33 and 44 unrelated individuals by genotyping 29 and 5 microsatellite markers on two chromosomes, respectively. Estimates of  $r^2$  (squared correlation coefficient) revealed significant LD only for loci < 1 cM apart. More recently, Du *et al.* [28] and McKay *et al.* [29] evaluated LD in pigs and cattle, respectively, by gen-

otyping syntenic single nucleotide polymorphisms (SNPs) and observing significant LD for marker pairs < 3 cM apart for pigs and 0.5 Mb for cattle. All of these results indicate that large numbers of markers from high-density maps are required to identify genes of interest using whole genome association studies in these species.

The USDA/ARS National Center for Cool and Cold Water Aquaculture (NCCCWA) has established a breeding program for rainbow trout, the use of molecular genetic technologies in this program is expected to enhance capabilities for selective breeding of important aquaculture production traits. To this end we have worked within international collaborations to develop genomic tools and technologies for rainbow trout [5,8,9,30-46] while initiating and characterizing our broodstock population with respect to genetic and phenotypic variation relevant to aquaculture production [47-57]. Our approaches include genetic linkage and LD mapping for the identification of QTL affecting traits of economic importance which will permit the development of marker/gene assisted selection strategies [58,59] and eventually genomic selection [60]. Characterizing the extent of LD in the NCCCWA rainbow trout broodstock population will support the use of integrated mapping approaches and facilitate the identification of genes affecting traits of interest by determining the marker densities required to conduct genome association scans. The NCCCWA broodstock population is closely related to commercial germplasm [49], therefore our findings also have the potential to impact industry selective breeding programs. In addition to LD we have evaluated effective population size ( $N_e$ ) in an effort to characterize the true breeding size of our population. This estimate of  $N_e$  should be considered when making decisions concerning selection pressure. To this end, we characterized extent of LD by genotyping 96 unrelated individuals with 49 markers spanning four chromosomes.

## Results

Genotypes were obtained for a total of 49 microsatellites on chromosomes OMY13 ( $n = 6$ ), OMY14 ( $n = 20$ ), OMY17 ( $n = 8$ ), and OMYSex ( $n = 15$ ). Genotyping success rate was 95.5% with 82 to 96 animals scored for each marker. The number of alleles per marker, marker heterozygosity, PIC, Allelic Diversity, and exact tests for departure from HWE proportions are reported in Table 1. The number of alleles per marker averaged 10.6 and ranged from 3 to 24. Marker heterozygosities ranged from 28.1% to 94.8% with an average of 69.8%. Polymorphism Information Content and Allelic Diversity ranged from 0.385 to 0.936 and 0.425 to 0.94, with averages of 0.745 and 0.773, respectively. A total of 19 loci had significant departure from HWE at  $P < 0.05$ , consisting of 3 markers

**Table 1: Exact test for Hardy-Weinberg equilibrium for microsatellite loci typed in 96 unrelated fish from the NCCCWA rainbow trout selective breeding program.**

Locus	GenBank Acc. or reference	OMY	Posit. Kos cM <sup>1</sup>	Number Alleles	PIC <sup>2</sup>	Heterozygosity	Allelic Diversity	$\chi^2$	DF	Pr > $\chi^2$	Exact P <sup>3</sup>
OMM3006	<a href="#">G73806</a>	13	39.9	13	0.844	0.802	0.858	127.6	78	0.0003	0.0615
OMM5092	<a href="#">CA348764</a>	13	39.9	5	0.705	0.708	0.746	20.7	10	0.0231	0.0152
OMY27DU	[4]	13	44.6	5	0.575	0.594	0.618	5.7	10	0.8402	0.7166
OMM1670	<a href="#">BV212162</a>	13	44.6	16	0.889	0.844	0.898	242.4	120	<.0001	0.0072
CA341677	<a href="#">CA341677</a>	13	44.6	6	0.544	0.604	0.590	99.5	15	<.0001	0.2147
OMM1687	<a href="#">BV212173</a>	13	45.7	3	0.390	0.417	0.501	7.2	3	0.0659	0.0325
Oneu8	<a href="#">U56708</a>	14	22.2	6	0.596	0.583	0.654	5.8	15	0.9824	0.8181
OMM1312	<a href="#">G73552</a>	14	49.2	16	0.882	0.885	0.891	125.5	120	0.3477	0.1418
OMYRGT2TUF	<a href="#">AB087587</a>	14	50.7	9	0.724	0.750	0.748	34.9	36	0.5218	0.6786
OMM1596	<a href="#">BV212112</a>	14	53.9	13	0.857	0.844	0.871	98.9	78	0.0552	0.0068
OMM5271	<a href="#">BV211986</a>	14	58.3	11	0.806	0.802	0.827	53.6	55	0.5286	0.4101
OmyRGT41TUF	[4]	14	58.3	16	0.876	0.854	0.885	114.9	120	0.6136	0.3824
Ogo1UW	<a href="#">AF007827</a>	14	58.3	5	0.620	0.635	0.679	98.9	10	<.0001	0.1218
OMM3089	<a href="#">BV718454</a>	14	62.5	4	0.612	0.760	0.679	6.1	6	0.4098	0.3408
OMM1447	<a href="#">BV079591</a>	14	63.4	22	0.936	0.823	0.940	237.6	231	0.3696	0.0017
OMM3115	<a href="#">AB162343</a>	14	63.4	12	0.843	0.490	0.859	219.2	66	<.0001	<.0001
BHMS429	<a href="#">AF256719</a>	14	64.3	12	0.890	0.885	0.898	50.5	66	0.9214	0.9010
OMYFGT5TUF	[4]	14	64.3	10	0.569	0.604	0.620	40.8	45	0.6515	0.3260
BHMS185	<a href="#">AF256675</a>	14	72.9	3	0.454	0.583	0.546	2.8	3	0.4239	0.2933
OMM1415	<a href="#">BV722101</a>	14	75.0	11	0.793	0.719	0.811	76.0	55	0.0317	0.0233
OMM1356	<a href="#">BV005148</a>	14	79.1	4	0.385	0.281	0.432	138.6	6	<.0001	<.0001
CR374305	<a href="#">CR374305</a>	14	81.2	13	0.790	0.646	0.809	439.9	78	<.0001	<.0001
OMM1467	<a href="#">BV079662</a>	14	82.7	12	0.866	0.906	0.878	79.2	66	0.1271	0.1546
OMM1643	<a href="#">BV212150</a>	14	86.5	11	0.764	0.750	0.785	89.7	55	0.0022	0.0354
OMM5143	<a href="#">BV211872</a>	14	106.2	11	0.828	0.427	0.846	271.4	55	<.0001	<.0001
OMM3044	<a href="#">BV722031</a>	14	130.0	8	0.739	0.396	0.775	242.4	28	<.0001	<.0001
OMM1712	<a href="#">BV212193</a>	17	81.7	13	0.855	0.823	0.868	67.9	78	0.7859	0.5445
CA367675	<a href="#">CA367675</a>	17	84.8	12	0.851	0.917	0.864	51.7	66	0.9017	0.8540
OMM5043	<a href="#">CA349167</a>	17	89.5	8	0.795	0.750	0.820	34.5	28	0.1836	0.1205

**Table 1: Exact test for Hardy-Weinberg equilibrium for microsatellite loci typed in 96 unrelated fish from the NCCCWA rainbow trout selective breeding program. (Continued)**

OMM3126	<u>BV683039</u>	17	90.3	18	0.904	0.833	0.911	144.1	153	0.6841	0.1027
OMM1437	<u>BV722116</u>	17	91.3	15	0.897	0.865	0.904	105.6	105	0.4648	0.1512
OMM1357	<u>BV005149</u>	17	105.4	17	0.867	0.833	0.877	207.1	136	<.0001	0.0305
OMM5227	<u>BX301679</u>	17	107.8	7	0.697	0.510	0.738	91.2	21	<.0001	<.0001
CA041953	<u>CA041953</u>	17	136.3	3	0.420	0.458	0.476	2.0	3	0.5711	0.4895
OMM1026	<u>AF346683</u>	sex	0.0	15	0.880	0.896	0.890	123.9	105	0.1007	0.5136
OMM1461	<u>BV079604</u>	sex	17.7	5	0.637	0.625	0.683	5.1	10	0.8870	0.7533
OMYRGT28TUF	<u>AB087599</u>	sex	21.1	15	0.830	0.760	0.846	172.6	105	<.0001	0.1674
OMM1000	<u>AF346664</u>	sex	21.1	3	0.385	0.417	0.425	0.5	3	0.9170	0.9396
OMM5031	<u>CA349143</u>	sex	29.7	9	0.722	0.865	0.753	59.8	36	0.0076	0.0005
OMM5032	<u>CA349143</u>	sex	29.7	11	0.842	0.844	0.857	94.7	55	0.0007	0.0895
BX076085	<u>BX076085</u>	sex	35.6	12	0.868	0.646	0.879	156.8	66	<.0001	<.0001
OMM1372	<u>BV005159</u>	sex	36.0	6	0.560	0.323	0.604	98.1	15	<.0001	<.0001
OMM1212	<u>BV722077</u>	sex	44.3	11	0.794	0.781	0.813	61.9	55	0.2421	0.2160
OMM1443	<u>BV079588</u>	sex	44.3	13	0.873	0.896	0.884	94.5	78	0.0984	0.0511
OMM3109	<u>BV718471</u>	sex	46.2	13	0.754	0.677	0.771	101.8	78	0.0363	0.0754
OMM1456	<u>BV079600</u>	sex	46.9	7	0.691	0.729	0.736	14.6	21	0.8447	0.7070
OMM1118	<u>AF375026</u>	sex	47.7	24	0.921	0.948	0.925	314.2	276	0.0564	0.1749
OMM1405	<u>BV722096</u>	sex	51.3	11	0.817	0.594	0.835	141.1	55	<.0001	<.0001
OMM1665	<u>BV212292</u>	sex	58.5	17	0.881	0.656	0.890	325.0	136	<.0001	<.0001

<sup>1</sup>Microsatellite loci position in the NCCCWA rainbow trout genetic map [9].

<sup>2</sup>Polymorphic information content (PIC).

<sup>3</sup>Exact *P*-value estimated using 10,000 permutations with SAS Proc Allele [70].

from OMY13, 9 from OMY14, 2 from OMY17, and 5 from OMYSex.

LD for syntenic loci was estimated for each individual chromosome and the four chromosomes combined (Figure 1), the genome-wide estimate for extent of significant LD ( $r^2 > 0.25$ ) was about 2 cM. The non-linear modeling of LD decline using the Sved (1971) equation indicates that the extent of significant LD decays sharply with physical distance. It also indicates that significant LD effects (i.e.,  $r^2 \geq 0.25$ ) are not expected beyond 5 cM.

The LD for non-syntenic loci was estimated for each chromosome pairing and all pairings of non-syntenic loci, the proportion of pairings having ( $r^2 > 0.25$ ) is shown in Figure 2. Clearly, the proportion of significant LD ( $r^2 \geq 0.25$ ) among non-syntenic loci was quite low

Effective population sizes were estimated for each chromosome and the four chromosomes as shown in Table 2. The estimated average effective population size was  $N_e = 145$  with an average  $N_e/N$  ratio of 0.45.

## Discussion

The extent of LD in the NCCCWA selective breeding program was evaluated by genotyping 49 microsatellites from four chromosomes on a total of 96 unrelated individuals from the 2005 ( $n = 43$ ) and 2006 ( $n = 53$ ) brood classes. As a result of an evolutionarily recent autopolyploid genome duplication event [61], many microsatellite markers for rainbow trout amplify two loci which can be scored independently and placed on genetic maps [4]. Although medium density genetic maps based on microsatellites including duplicated markers have been constructed [6,9], they are problematic for use in population genetic analysis such as whole genome association studies as alleles from the two loci often have overlapping and/or identical allele sizes. Thus the number of loci available for use in whole genome association studies is much less than what is available for segregation analyses. Only evaluating single locus markers resulted in observation of a 6 cM region of OMY13 with 6 loci, a 108 cM region of OMY14 with 20 loci, a 55 cM region of OMY17 with 8 loci, and a 58 cM region of OMYSex with 15 loci. Overall, 227 cM of the 2927 cM genome was represented in this study. Of interest is the 55 cM region of OMY14 spanning the region from 75-130 cM where 6/7 loci depart from HWE. It is possible that these loci are under selection in the NCCCWA broodstock population which includes selective pressure for disease resistance and growth.

The level of LD between syntenic loci decayed rapidly at distances greater than 2 cM which is similar to observations of LD in other agriculturally important species including cattle [20], sheep [25], pigs [27,28] and chick-

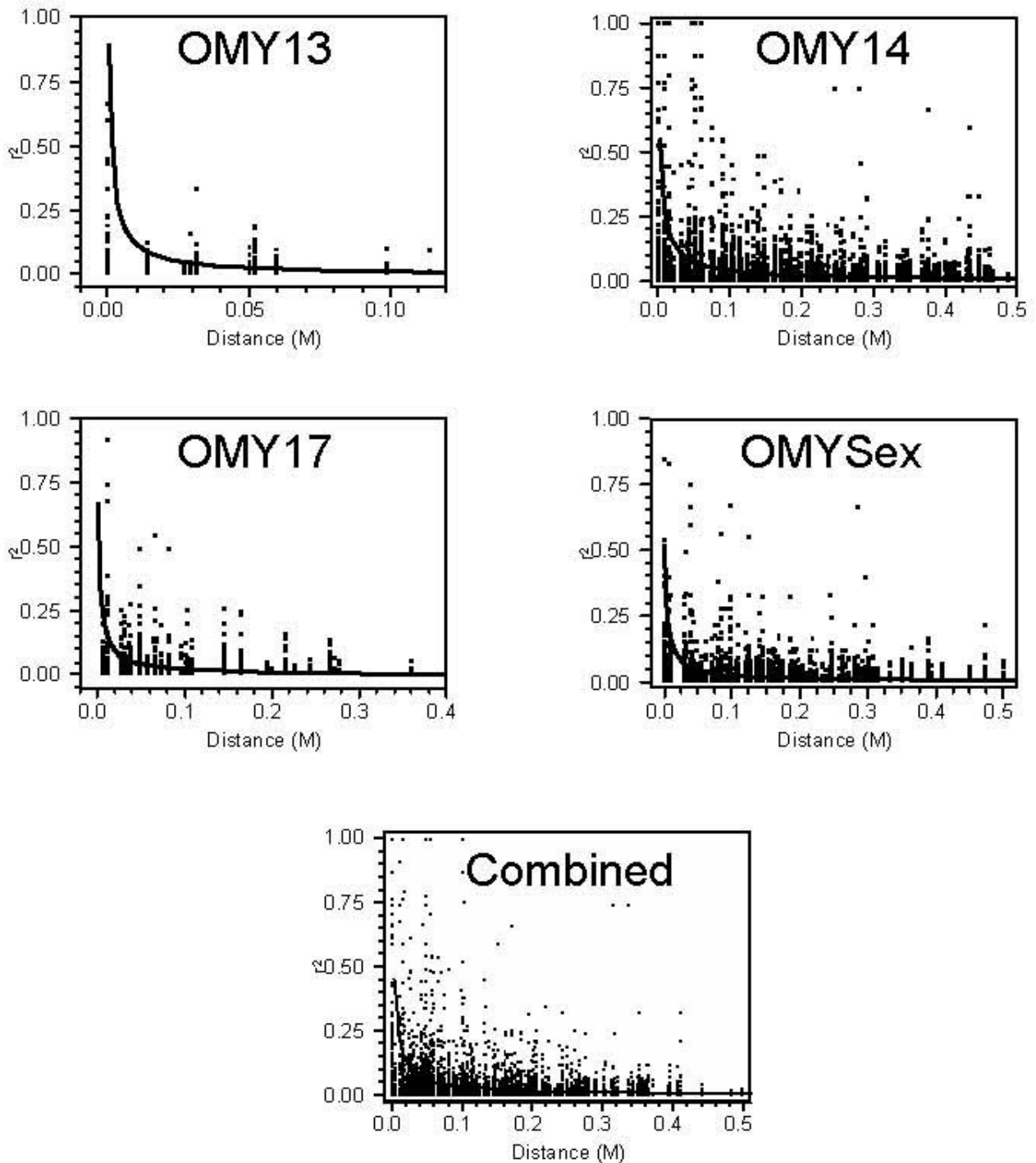
ens [26]. However, significant LD was also observed up to 40 cM as reported by Farnir *et al.* [20] and McRae *et al.* [25] (Figure 1). The average proportion of significant LD ( $r^2 > 0.25$ ) between non-syntenic loci was under .005, however OMY13 and OMY17, which have a common ancestor chromosome resulting from the salmonid whole genome duplication event, showed a significantly higher proportion of significant LD.

We acknowledge that the LD range reported here may be up-biased because evidences for structure and admixture have been shown in the NCCCWA Broodstock population [62] and we used unrelated individuals from 2005 and 2006 brood years that are being improved for disease resistance and production traits, respectively, in the NCCCWA selective breeding program.

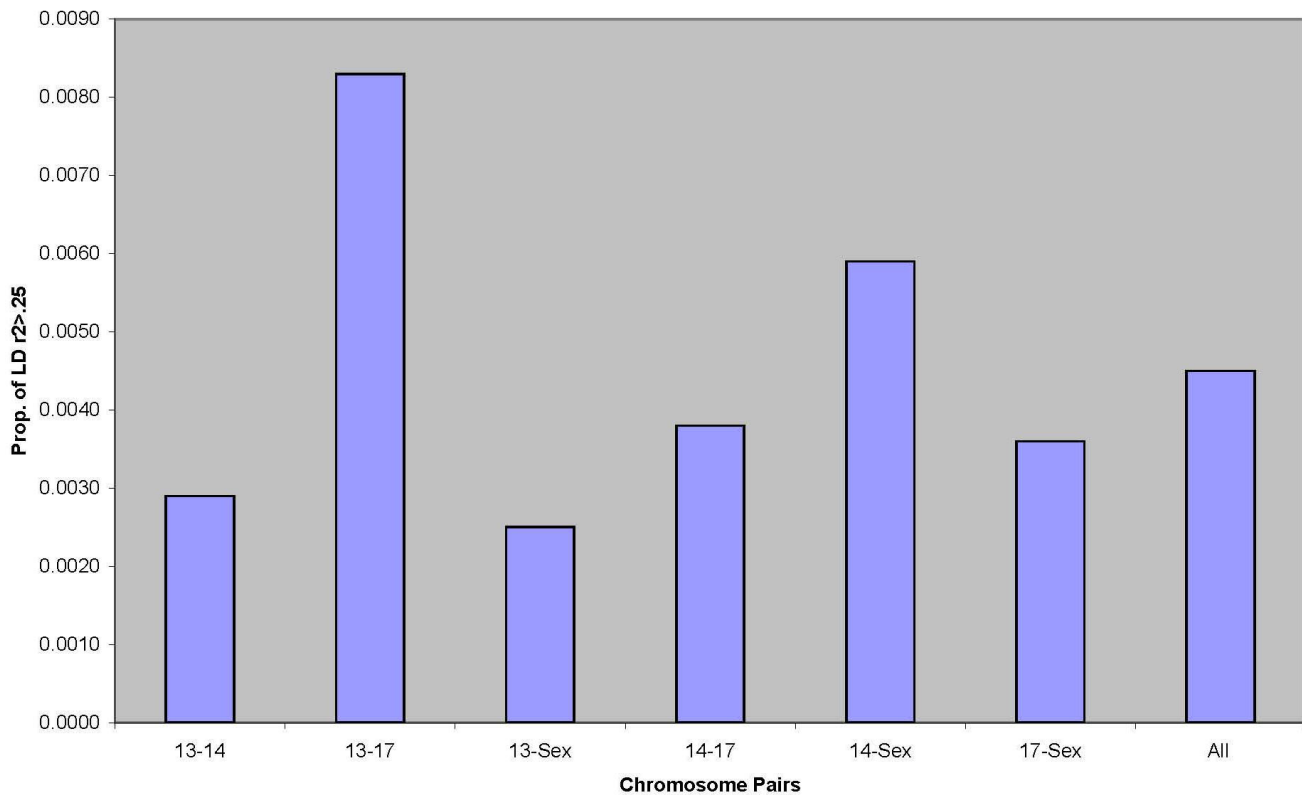
As the NCCCWA selective breeding program continues throughout successive generations, we expect that the effective population size will continue to decrease. Our estimate of  $N_e = 145$  (Table 2) based on genome wide LD indicates that this population will respond well to high selection intensity as it has for a single generation of breeding for resistance to the causative agent of bacterial cold-water disease [50]. Using this estimate we can observe the effects of selection and identify correlations with the effects of inbreeding, usually first observed in reproductive success. However, the range of effective population size based on individual chromosomes was 75.51 - 203.35, possibly indicating that suites of genes on each chromosome are disproportionately under selection pressure. This also establishes the need to develop whole genome applications for these species for studies of LD. Although  $F$ -statistics have been used to estimate population genetic parameters [49], characterization of LD will also provide valuable information about population substructure and should be a factor in developing strategies aimed at the identification of associations between markers and QTL.

We chose to evaluate the extent of LD by genotyping representatives of each year class a single generation after the population was founded by mixing germplasm from multiple sources. Our resulting characterization of LD not only provides information about population structure, but serves to document the degree of genetic diversity used to initiate the broodstock program. We expect our estimates of  $N_e$  to be higher than populations with limited genetic diversity and no introduction of new germplasm in recent past generations.

The estimated  $N_e/N$  ratio of 0.45 in this study is between the ranges of estimates reported in this species in previous studies [63-65]. Other studies in salmonids consistently



**Figure 1**  
**Decline of linkage disequilibrium ( $r^2$ ) with distance (recombination rate in Morgans) for chromosomes 13, 14, 17 and Sex.** The estimates of  $r^2$  for pairs of markers were adjusted for experimental sample size  $1/n$ , where  $n$  is the chromosome sample size ( $n = 192$ ). The predicted LD value plot (filled non-linear curve) was estimated fitting the equation  $LD_{ij} = 1 / (1 + kb_j d_{ij}) + e_{ij}$  performing non-linear modeling with JMP® Genomics 3.1 (SAS Institute Inc., Carey, NC, 2007). Here,  $LD_{ij}$  is the observed LD for marker pair  $i$  in chromosome  $j$ ,  $d_{ij}$  is the distance in Morgans for marker pair  $i$  in chromosome  $j$ ,  $b_j$  is the estimate of effective population size for chromosome  $j$ , and the constant  $k = 2$  for sex chromosome and  $k = 4$  for autosomes.



**Figure 2**  
**Proportion of markers pairs with significant extent of LD ( $r^2 \geq 0.25$ ).** All marker pairs were evaluated in addition to pairwise combinations of all non-syntenic loci.

reported that the variance in reproductive success is the key factor to reduce  $N_e/N$  in salmon populations [64,66].

Characterizing the extent and distribution of LD helps to determine the required marker density for LD mapping and genomic selection as they both require markers to be in LD with QTL. Our observation of significant syntenic LD at distances over 2 cM has implications for designing genome wide association studies in this population and on this species. The sex averaged map is roughly 3000 cM long; therefore 1500 markers are required to identify loci of interest. The male map having 2500 cM would require 1125 markers; the female map having 4300 cM would require 2150 markers. Currently about 1800 microsatellite markers are available for genome analyses in trout. However, we must take into consideration that: 1) roughly 33% are duplicated; 2) on average 70% are informative based on estimates of heterozygosity; and 3) these markers are not necessarily spaced at regular intervals throughout the genome.

## Conclusions

To effectively conduct whole genome association studies the number of available markers for rainbow trout must

be increased. Whereas genotyping microsatellites can be very expensive and time consuming process, it is likely that a marker system that enables high-throughput genotyping protocols such as single nucleotide polymorphisms will be the basis for LD mapping and genomic selection studies in the near future.

## Methods

### Germplasm

The NCCCWA rainbow trout selective breeding program was initiated in 2002 through 2004 with the acquisition of fish from Troutlodge, Inc. (Sumner, WA), the Donaldson strain from the University of Washington (Seattle, WA), the House Creek strain from the College of Southern Idaho (Twin Falls, ID), and the Shasta strain from the Ennis National Fish Hatchery, (Ennis, MT) [49]. To date, selective breeding for increased aquaculture production efficiency has been conducted by evaluating and selecting for growth traits on even years and resistance to challenge by the bacterial pathogen *Flavobacterium psychrophilum* on odd years [47,50]. To identify individuals that could best represent the genetic variation contained within this broodstock population, we identified 96 unrelated individuals (no siblings or half-siblings) from the 2005 and

**Table 2: Effective population size ( $N_e$ ) estimated from linkage disequilibrium<sup>1</sup> by fitting nonlinear regression model<sup>2</sup> in a rainbow trout broodstock population<sup>3</sup>.**

Chromosome	Number of marker pairs	Average intermarker distance (cM)	Genome coverage (cM)	$\hat{N}_e$	Standard error	Confidence interval		$\hat{N}_e / N^4$
						-95%	+95%	
13	421	1.16	5.80	178.56	75.38	65.77	448.56	0.56
14	8817	5.67	107.80	75.51	5.62	65.24	87.14	0.24
17	1499	7.80	54.60	122.08	19.16	89.45	162.23	0.38
sex	5813	4.18	58.50	203.35	16.94	170.92	240.54	0.64
<b>Total</b>	16550.00	18.81	226.70	579.50	117.10	391.38	938.47	1.82
<b>Mean</b>	4137.50	4.70	56.68	144.88	29.27	97.85	234.62	0.45
<b>SD</b>				57.40	31.30	50.01	155.77	0.18

<sup>1</sup>Pairwise Linkage disequilibrium (LD) was estimated using the ALLELE procedure of the software package SAS<sup>®</sup>, version 9.3.1 (SAS Institute 2007).

<sup>2</sup>The nonlinear regression model was fitted using JMP<sup>®</sup> Genomics 3.1 (SAS Institute Inc., Carey, NC, 2007).

$$y_i = \frac{1}{(\alpha + \beta c_i)} + e_i$$

Where  $y_i = (r^2 - 1/n)$  is LD measure adjusted for chromosome sample size  $n$ , for marker pair  $i$  at recombination rate  $c_i$  (in Morgans). The constant  $k$  had values of  $k = 2$  for sex chromosome and  $k = 4$  for autosomes. The  $c_i$ 's were estimates of recombination rate from two-point linkage analysis [9]. First, the  $e_i$  residuals were estimated by non-linear fitting of the above model with JMP<sup>®</sup> Genomics 3.1 (SAS Institute Inc., Carey, NC,

2007). Then, the parameters  $\alpha$  and  $\beta$  were estimated iteratively by least squares; in this model  $\hat{N}_e = \hat{\beta} / k$ .

<sup>3</sup>Unrelated individuals ( $n = 96$ ) representing the 2005/2006 brood classes were genotyped with 49 microsatellite markers.

<sup>4</sup>Number of potential breeders ( $N = 320$ ).

2006 brood year classes ( $n = 43$  and  $53$ , respectively) to represent the 144 and 177 fish that were actually used to produce select matings. The 2005 year class, part of our odd year selection for disease resistance, is the first generation after the 2003 founder population. The 2006 year class, part of our even year selection for growth, is the second generation after the founder population in 2002. However, a significant contribution of new germplasm was also introduced in 2004. Fin clips from anesthetized fish were collected for DNA extraction as outlined in a protocol approved by the NCCCWA IACUC, number 025-1-26-05. The extraction protocols followed the phenol-chloroform method described in Sambrook and Russell [67]. DNA samples were quantified by spectrophotometer (Beckman DU 640, Beckman Instruments, St. Louis, MO, USA) and diluted to a concentration of 12.5 ng/ul for PCR.

### Genotyping

Microsatellite markers from four chromosomes ( $n = 49$ ) were selected from the NCCCWA genetic map [9] for genotyping the DNA panel of 96 fish representing NCCCWA broodstock (Table 1). This included markers mapped to the sex chromosome (OMYSex), chromosome 14 having the largest linkage group in cM (OMY14), and paralogous

regions of chromosomes 13 and 17 (OMY13, OMY17) as defined by duplicated loci resulting from an evolutionarily recent genome duplication event [4,9,61]. Only single locus markers were utilized. Markers were either genotyped using the tailed protocol [68] or by direct fluorescent labeling (with FAM, HEX, or NED) of the forward primer according to manufacturer protocols [[69] USA]. Primer pairs were obtained from commercial sources (forward primers labeled with FAM or HEX from Alpha DNA, Montreal, Quebec, Canada, or NED from ABI, Foster City, CA, USA). PCR reactions consisted of 12  $\mu$ l reaction volumes containing 12.5 ng DNA, 1.5-2.5 mM MgCl<sub>2</sub>, 1.0  $\mu$ M of each primer, 200  $\mu$ M of dNTPs, 1 $\times$  manufacturer's reaction buffer and 0.5 units Taq DNA polymerase. Thermal cycling consisted of an initial denaturation at 95°C for 15 min followed by 30 cycles of 95°C for 1 min, annealing temperature for 45 s, 72°C extension for 45 s and a final extension at 72°C for 10 min. PCR products were visualized on agarose gels after staining with ethidium bromide. Markers were grouped in combinations of two or three markers based on differences in fluorescent dye color and amplicon size. Three  $\mu$ l of each PCR product was added to 20  $\mu$ l of water, 1  $\mu$ l of the diluted sample was added to 12.5  $\mu$ l of loading mixture made up with 12  $\mu$ l of HiDi formamide and 0.5 of Genscan 400 ROX internal



size standard. Samples were denatured at 95°C for 5 min and kept on ice until loading on an automated DNA sequencer ABI 3730 DNA Analyzer (ABI, Foster City, CA, USA). Output files were analyzed using GeneMapper version 3.7 (ABI, Foster City, CA, USA), formatted using Microsoft Excel and stored in Microsoft Access database.

### Analysis

The initial dataset for analysis included marker genotype data for 49 microsatellite loci typed on 96 unrelated individuals. These marker genotype data were analyzed to estimate frequency of alleles per marker using the ALLELE procedure of software package SAS®, version 9.3.1 [70]. Within each marker, alleles with frequency  $\leq 0.01$  were merged to minimize the up biasing effect of rare alleles on LD estimates. Then, the marker alleles initially recorded in size of fragments were recoded into a consecutive-numbered allele system using the computer program RECODE [71]. This dataset with recoded marker genotypes were divided into files including marker loci corresponding to each of the four linkage groups analyzed in this study. Subsequently, these recoded marker genotype data were used in haplotype reconstruction and linkage disequilibrium analysis.

### Haplotype reconstruction

For each linkage group, the most likely haplotype configuration for each individual was estimated using the software package PHASE, version 2.1, following procedures described by [72] and [73]. Briefly, this is a statistical method for inferring haplotypes from unphased genotype data for a sample of "unrelated" individuals from a population; population haplotype frequencies are assumed under Hardy-Weinberg equilibrium (HWE) but PHASE has proven robust to deviations from HWE, the effect of population structure and moderate amounts of recombination. The estimated haplotype data was formatted for subsequent analysis with the ALLELE procedure of the software package SAS®, version 9.3.1 [70]. The haplotype data was formatted into SAS ALLELE procedure format (option TALL for data input format) using a Java script (Haplo\_2SAS.jar) written by Christopher Schmitt. This script is available for academic use upon request to RLV. This formatted dataset was used in linkage disequilibrium analysis.

### Hardy-Weinberg equilibrium

The HWE analysis for each of the 49 loci typed in 96 unrelated individuals was performed using the ALLELE procedure of the software package SAS®, version 9.3.1 [70] using the option TALL for data input format. The input data were the most likely haplotypes reconstructed for each individual using the computer program PHASE and reformatted into SAS format as outlined above. The exact test for HWE for each microsatellite loci (exact *P*-value) was estimated using 10,000 permutations.

### Linkage disequilibrium between syntenic loci

The LD analysis was performed between marker pairs within each linkage group using the ALLELE procedure of software package SAS®, version 9.3.1 [70]. We used the option TALL for data input format, and the HAPLO = GIVEN option which indicates that the haplotypes have been observed, and thus the observed haplotype frequencies were used in the LD test statistic and measures.

The linkage (or gametic) disequilibrium coefficient  $D_{uv}$  between two alleles from marker loci *M* and *N*, respectively, was estimated using this expression [74],

$$D_{uv} = p_{uv} - p_u p_v$$

Where  $p_{uv}$  is probability that an individual receives the haplotype  $M_u N_v$  for marker loci *M* and *N*,  $p_u$  is the probability of the *u* allele, and  $p_v$  is the probability of the *v* allele.

The ALLELE procedure calculates the maximum likelihood estimates,  $\tilde{D}_{uv}$ , of the LD coefficient between a pair of alleles at different markers. This procedure calculates an overall chi-square statistic to test that all of the  $D_{uv}$ 's between two markers are zero as follows,

$$\chi^2_T = \sum_{u=1}^k \sum_{v=1}^l \frac{(2n) \tilde{D}_{uv}^2}{\tilde{p}_u \tilde{p}_v}$$

Which has  $(k-1)$  and  $(l-1)$  degrees of freedom for markers with *k* and *l* alleles, respectively. A Monte Carlo estimate of the exact *P*-value for testing the hypothesis was calculated by conditioning on the haplotype counts. The significance level is obtained by permuting the alleles at one locus to form  $2^n$  new two-locus haplotypes. We used 10,000 permutations to estimate the exact *P*-values.

We used the squared correlation coefficient ( $r^2$ ) as the LD measure for each pair of alleles  $M_u$  and  $N_v$  located at loci *M* and *N*, respectively [75],

$$r^2 = \frac{D^2}{p_1 p_2 q_1 q_2}$$

Where  $D = p_{11} p_{12} - p_{12} p_{21}$  is the LD coefficient, which can be directly estimated using the observed haplotype frequencies  $\tilde{p}_{uv}$  when using the option HAPLO = GIVEN with the ALLELE procedure of SAS®, version 9.3.1 [70]. Since these measures are designed for biallelic markers, the measures are calculated for each allele at locus *M* with each allele at locus *N*, where all other alleles at each locus are combined to represent one allele. Thus for each allele

$M_u$  in turn,  $\tilde{p}_1$  is used as the frequency of allele  $M_u$ , and  $\tilde{p}_2$  represents the frequency of "not  $M_u$ "; similarly for each  $N_v$  in turn,  $\tilde{q}_1$  represents the frequency of allele  $N_v$ , and  $\tilde{q}_2$  the frequency of "not  $N_v$ ."

#### Linkage disequilibrium decay with distance

The decline of linkage disequilibrium with distance (recombination rate in Morgans) was estimated by fitting the following equation [76],

$$r_{ij} = \frac{1}{(1+kb_j d_{ij})} + e_{ij}$$

Where  $r_{ij} = (r^2 - 1/n)$  is the observed LD for marker pair  $i$  in chromosome  $j$ , the constant  $k = 2$  for sex chromosome and  $k = 4$  for autosomes,  $d_{ij}$  is the recombination rate from two-point linkage analysis for marker pair  $i$  in chromosome  $j$ ,  $b_j$  is the estimate of effective population size for chromosome  $j$ , and  $e_{ij}$  is a random residual. The estimates of  $r^2$  for pairs of markers were adjusted for experimental sample size  $1/n$ , where  $n$  is the chromosome sample size ( $n = 192$ ). We performed the non-linear modeling with JMP® Genomics 3.1 (SAS Institute Inc., Carey, NC, 2007).

#### Linkage disequilibrium between nonsyntenic loci

We estimated LD adjusted for experimental sample size ( $r^2 - 1/n$ ) between pairs of non-syntenic loci using the ALLELE procedure of software package SAS®, version 9.3.1 [70]. As input data, we used marker genotype data with alleles recoded into a consecutive-numbered system (non-phased marker genotype data). In the analysis, we used the option HAPLO = EST which indicates that the maximum likelihood estimates of the haplotype frequencies are used to calculate the LD test statistic as well as the LD measures. For LD estimation among non-syntenic loci, we used the ALLELE procedure options ALLELEMIN = GENOMIN = HAPLOMIN = 0.01. This last statement ensures that only alleles, genotypes, and haplotypes with frequency  $\geq 0.01$  are used in the LD analysis.

#### Effective population size

The analysis was based on the known relationship between LD as measured by  $r^2$  (squared correlation of allele frequencies at a pair of loci) and effective population size  $N_e$ ,

$$E(r^2) = \frac{1}{(\alpha + kN_e c)} + 1/n$$

Where  $c$  is the recombination rate between the microsatellite loci and  $n$  is the experimental sample size. The constant  $\alpha = 1$  in the absence of mutation [76] and  $\alpha = 2$  if mutation is taken into account [77-79]. The constant  $k$  was set to  $k = 2$  for sex chromosome and  $k = 4$  for autosomes.

Given the formulae described in the linkage disequilibrium section, and knowing  $r^2$  and  $c$ , we estimated  $N_e$  for each chromosome by fitting this nonlinear regression model,

$$r_{ij} = \frac{1}{(\alpha_j + \beta_j c_{ij})} + e_{ij}$$

Where  $r_{ij} = (r^2 - 1/n)$  is the observed LD (adjusted for chromosome sample size  $n$ ) for marker pair  $i$  in chromosome  $j$ ,  $c_{ij}$  is the recombination rate from two-point linkage analysis for marker pair  $i$  in chromosome  $j$ . The parameter  $\beta_j$  is the estimator of effective population size for chromosome  $j$  where  $\hat{N}_e = \beta_j / k$ . The parameters  $\alpha_j$  and  $\beta_j$  were estimated iteratively using non-linear modeling with JMP® Genomics 3.1 (SAS Institute Inc., Carey, NC, 2007).

#### Authors' contributions

CR participated in design, marker selection, genotyping, and drafted the manuscript, RV participated in design, performed the calculations for LD and effective population size. Both authors read and approved the final manuscript.

#### Acknowledgements

We are grateful to Roseanna Long and Kristy Shewbridge for contributing their technical expertise to this project.

#### References

1. Liu ZJ, Ed: **Aquaculture Genome Technologies**. Ames, Iowa: Blackwell Publishing; 2007.
2. Thorgaard GH, Bailey GS, Williams D, Buhler DR, Kaattari SL, Ristow SS, Hansen JD, Winton JR, Bartholomew JL, Nagler JJ, et al.: **Status and opportunities for genomics research with rainbow trout**. *Comp Biochem Physiol B Biochem Mol Biol* 2002, **133**:609-646.
3. Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH: **A detailed linkage map of rainbow trout produced using doubled haploids**. *Genetics* 1998, **148**:839-850.
4. Sakamoto T, Danzmann RG, Gharbi K, Howard P, Ozaki A, Khoo SK, Woram RA, Okamoto N, Ferguson MM, Holm LE, et al.: **A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates**. *Genetics* 2000, **155**:1331-1345.
5. Nichols KM, Young WP, Danzmann RG, Robison BD, Rexroad CE, Noakes M, Phillips RB, Bentzen P, Spies I, Knudsen K, et al.: **A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*)**. *Anim Genet* 2003, **34**:102-115.
6. Guyonard R, Mauger S, Tabet-Canale K, Martineau S, Genet C, Krieg F, Quillet E: **A type I and type II microsatellite linkage map of**

- rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms. *BMC Genomics* 2006, **7**:302.
7. Phillips RB, Nichols KM, DeKoning JJ, Morasch MR, Keatley KA, Rexroad C, Gahr SA, Danzmann RG, Drew RE, Thorgaard GH: **Assignment of rainbow trout linkage groups to specific chromosomes.** *Genetics* 2006, **174**:1661-1670.
  8. Danzmann RG, Cairney M, Davidson WS, Ferguson MM, Gharbi K, Guyomard R, Holm LE, Leder E, Okamoto N, Ozaki A, et al.: **A comparative analysis of the rainbow trout genome with 2 other species of fish (Arctic charr and Atlantic salmon) within the tetraploid derivative Salmonidae family (subfamily: Salmoninae).** *Genome* 2005, **48**:1037-1051.
  9. Rexroad CE, Palti Y, Gahr SA, Vallejo RL: **A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*).** *BMC Genet* 2008, **9**:74.
  10. Leder EH, Danzmann RG, Ferguson MM: **The candidate gene, Clock, localizes to a strong spawning time quantitative trait locus region in rainbow trout.** *J Hered* 2006, **97**:74-80.
  11. Martinez V, Thorgaard G, Robison B, Sillanpaa MJ: **An application of Bayesian QTL mapping to early development in double haploid lines of rainbow trout including environmental effects.** *Genet Res* 2005, **86**:209-221.
  12. Nichols KM, Broman KW, Sundin K, Young JM, Wheeler PA, Thorgaard GH: **Quantitative trait loci x maternal cytoplasmic environment interaction for development rate in *Oncorhynchus mykiss*.** *Genetics* 2007, **175**:335-347.
  13. O'Malley KG, Sakamoto T, Danzmann RG, Ferguson MM: **Quantitative trait loci for spawning date and body weight in rainbow trout: testing for conserved effects across ancestrally duplicated chromosomes.** *J Hered* 2003, **94**:273-284.
  14. Perry GM, Danzmann RG, Ferguson MM, Gibson JP: **Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*).** *Heredity* 2001, **86**:333-341.
  15. Perry GM, Ferguson MM, Sakamoto T, Danzmann RG: **Sex-linked quantitative trait loci for thermotolerance and length in the rainbow trout.** *J Hered* 2005, **96**:97-107.
  16. Robison BD, Wheeler PA, Sundin K, Sikka P, Thorgaard GH: **Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*).** *J Hered* 2001, **92**:16-22.
  17. Zimmerman AM, Evenhuis JP, Thorgaard GH, Ristow SS: **A single major chromosomal region controls natural killer cell-like activity in rainbow trout.** *Immunogenetics* 2004, **55**:825-835.
  18. Ozaki A, Sakamoto T, Khoo S, Nakamura K, Coimbra MR, Akutsu T, Okamoto N: **Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*).** *Mol Genet Genomics* 2001, **265**:23-31.
  19. Nakamura K, Ozaki A, Akutsu T, Iwai K, Sakamoto T, Yoshizaki G, Okamoto N: **Genetic mapping of the dominant albino locus in rainbow trout (*Oncorhynchus mykiss*).** *Mol Genet Genomics* 2001, **265**:687-693.
  20. Farnir F, Coppieters W, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Mni M, et al.: **Extensive genome-wide linkage disequilibrium in cattle.** *Genome Res* 2000, **10**:220-227.
  21. Kruglyak L: **Prospects for whole-genome linkage disequilibrium mapping of common disease genes.** *Nat Genet* 1999, **22**:139-144.
  22. Meuwissen TH, Karlsen A, Lien S, Olsaker I, Goddard ME: **Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping.** *Genetics* 2002, **161**:373-379.
  23. Lewontin RC: **The detection of linkage disequilibrium in molecular sequence data.** *Genetics* 1995, **140**:377-388.
  24. Vallejo RL, Li YL, Rogers GW, Ashwell MS: **Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population.** *J Dairy Sci* 2003, **86**:4137-4147.
  25. McRae AF, McEwan JC, Dodds KG, Wilson T, Crawford AM, Slate J: **Linkage disequilibrium in domestic sheep.** *Genetics* 2002, **160**:1113-1122.
  26. Heifetz EM, Fulton JE, O'Sullivan N, Zhao H, Dekkers JC, Soller M: **Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations.** *Genetics* 2005, **171**:1173-1181.
  27. Harmegnies N, Farnir F, Davin F, Buys N, Georges M, Coppieters W: **Measuring the extent of linkage disequilibrium in commercial pig populations.** *Anim Genet* 2006, **37**:225-231.
  28. Du FX, Clutter AC, Lohuis MM: **Characterizing linkage disequilibrium in pig populations.** *Int J Biol Sci* 2007, **3**:166-178.
  29. McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppieters W, Crews D, Dias Neto E, Gill CA, Gao C, et al.: **Whole genome linkage disequilibrium maps in cattle.** *BMC Genet* 2007, **8**:74.
  30. Coulbaly I, Gharbi K, Danzmann RG, Yao J, Rexroad CE: **Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags.** *Anim Genet* 2005, **36**:309-315.
  31. Gahr SA, Rexroad CE, Rise ML, Hunt P, Koop B: **A survey of expressed sequence tags from the rainbow trout (*Oncorhynchus mykiss*) pituitary.** *Anim Biotechnol* 2007, **18**:213-230.
  32. Koop BF, Schalburg KRv, Leong J, Walker N, Lieph R, Cooper GA, Robb A, Beetz-Sargent M, Holt RA, Moore R, et al.: **A salmonid EST genomic study: genes, duplications, phylogeny and microarrays.** *BMC Genomics* 2008, **9**:545.
  33. Palti Y, Danzmann RG, Rexroad CE: **Characterization and mapping of 19 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*).** *Anim Genet* 2003, **34**:153-156.
  34. Palti Y, Fincham MR, Rexroad CE: **Characterization of 38 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*).** *Molecular Ecology Notes* 2002, **2**:449-452.
  35. Palti Y, Gahr SA, Hansen JD, Rexroad CE: **Characterization of a new BAC library for rainbow trout: evidence for multi-locus duplication.** *Anim Genet* 2004, **35**:130-133.
  36. Ramachandra RK, Salem M, Gahr S, Rexroad CE, Yao J: **Cloning and characterization of microRNAs present during early embryonic development in rainbow trout (*Oncorhynchus mykiss*).** *BMC Developmental Biology* 2008, **8**:41.
  37. Rexroad CE, Coleman RL, Gustafson AL, Hershberger WK, Killefer J: **Development of rainbow trout microsatellite markers from repeat enriched libraries.** *Mar Biotechnol (NY)* 2002, **4**:12-16.
  38. Rexroad CE, Coleman RL, Hershberger WK, Killefer J: **Rapid communication: Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout.** *J Anim Sci* 2002, **80**:541-542.
  39. Rexroad CE, Coleman RL, Hershberger WK, Killefer J: **Eighteen polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*).** *Anim Genet* 2002, **33**:76-78.
  40. Rexroad CE, Coleman RL, Martin AM, Hershberger WK, Killefer J: **Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*).** *Anim Genet* 2001, **32**:317-319.
  41. Rexroad CE, Lee Y, Keele JW, Karamycheva S, Brown G, Koop B, Gahr SA, Palti Y, Quackenbush J: **Sequence analysis of a rainbow trout cDNA library and creation of a gene index.** *Cytogenet Genome Res* 2003, **102**:347-354.
  42. Rexroad CE, Palti Y: **Development of Ninety-Seven Polymorphic Microsatellite Markers for rainbow trout (*Oncorhynchus mykiss*).** *Transactions of the American Fisheries Society* 2003, **132**:1214-1221.
  43. Rexroad CE, Rodriguez MF, Coulbaly I, Gharbi K, Danzmann RG, DeKoning J, Phillips R, Palti Y: **Comparative mapping of expressed sequence tags containing microsatellites in rainbow trout (*Oncorhynchus mykiss*).** *BMC Genomics* 2005, **6**:54.
  44. Rodriguez MF, Gahr SA, Rexroad CE, Palti Y: **Rapid microsatellite detection from rainbow trout (*Oncorhynchus mykiss*) bacterial artificial chromosome using PCR screening method.** *Marine Biotechnology* 2006, **8**:346-350.
  45. Salem M, Kenney B, Rexroad CE, Yao J: **Development of a 37K High-density oligo-nucleotide microarray for rainbow trout.** *Journal of Fish Biology* 2007, **72**:2187-2206.
  46. Phillips RB, Nichols KM, DeKoning JJ, Morash MR, Keatley KA, Rexroad CE, Danzmann RG, Drew RE, Thorgaard GH: **Assignment of Rainbow Trout Linkage Groups to Specific Chromosomes.** *Genetics* 2006, **174**:1661-1670.
  47. Lankford S, Weber G: **Associations between plasma growth hormone, insulin-like growth factor-I and cortisol, with stress responsiveness and growth performance in a selective breeding program for rainbow trout (*Oncorhynchus mykiss*).** *North American Journal of Aquaculture* 2006, **68**:151-159.

48. Silverstein J: **Using genetic variation to understand control of feed intake in fish.** *Fish Physiology and Biochemistry* 2004, **27**:173-178.
49. Silverstein JT, King T, Rexroad CE: **Genetic Variation Measured by Microsatellites Among Three Strains of Domesticated Rainbow Trout.** *Aquaculture Research* 2004, **35**:40-48.
50. Silverstein JT, Vallejo R, Palti Y, Leeds T, Rexroad CE, Welch T, Wiens G, Ducrocq V: **Rainbow trout resistance to bacterial cold-water disease is moderately heritable and is not adversely correlated with growth.** *Journal of Animal Science* 2009, **87**:860-867.
51. Vallejo RL, Rexroad CE, Silverstein JT, Janss LL, Weber GM: **Evidence of major genes affecting stress response in rainbow trout using Bayesian methods of complex segregation analysis.** *J Anim Sci* 2009, **87**:3490-3505.
52. Silverstein J, Hostuttler M, Blemings K: **Strain differences in feed efficiency measured as residual feed intake in rainbow trout.** *Aquaculture Research* 2005, **36**:704-711.
53. Hershberger W, Hostuttler M: **Variation in Time to First Cleavage in Rainbow Trout *Oncorhynchus mykiss* Embryos: A Major Factor in Induction of Tetraploids.** *Journal of the World Aquaculture Society* 2005, **36**:96-102.
54. Silverstein J: **Relationships among feed intake, feed efficiency and growth in juvenile rainbow trout (*Oncorhynchus mykiss*).** *North American Journal of Aquaculture* 2006, **68**:168-175.
55. Weber G, Silverstein J: **Evaluation of a stress response for use in a selective breeding program for improved growth and disease resistance in rainbow trout (*Oncorhynchus mykiss*).** *North American Journal of Aquaculture* 2006, **69**:69-79.
56. Weber G, Vallejo R, Lankford S, Silverstein J, Welch T: **Cortisol Response to a Crowding Stress: Heritability and Association with Disease Resistance to *Yersinia ruckeri* in Rainbow Trout.** *North American Journal of Aquaculture* 2008, **70**:425-433.
57. Hadidi S, Glenney G, Welch T, Silverstein J, Wiens G: **Spleen Size Predicts Resistance of Rainbow Trout to *Flavobacterium psychrophilum* Challenge.** *Journal of Immunology* 2008, **180**:4156-4165.
58. Dekkers JC: **Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons.** *J Anim Sci* 2004, **82**(E-Suppl):E313-328.
59. Soller M, Beckman J: **Genetic polymorphism in varietal identification and genetic improvement.** *Theoretical and Applied Genetics* 1983, **67**:25-33.
60. Meuwissen TH, Hayes BJ, Goddard ME: **Prediction of total genetic value using genome-wide dense marker maps.** *Genetics* 2001, **157**:1819-1829.
61. Allendorf FW, Thorgaard GH: **Tetraploidy and the evolution of salmonid fishes.** In *Evolutionary Genetics of Fishes* Edited by: Turner BJ. New York: Plenum Press; 1984:1-46.
62. Johnson NA, Rexroad CE, Hallerman EM, Vallejo RL, Palti Y: **Development and evaluation of a new microsatellite multiplex system for parental allocation and management of rainbow trout (*Oncorhynchus mykiss*) broodstocks.** *Aquaculture* 2007, **266**:53-62.
63. Araki H, Waples RS, Ardren WR, Cooper B, Blouin MS: **Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms.** *Mol Ecol* 2007, **16**:953-966.
64. Ardren WR, Kapuscinski AR: **Demographic and genetic estimates of effective population size ( $N_e$ ) reveals genetic compensation in steelhead trout.** *Mol Ecol* 2003, **12**:35-49.
65. Heath DD, Busch C, Kelly J, Atagi DY: **Temporal change in genetic structure and effective population size in steelhead trout (*Oncorhynchus mykiss*).** *Mol Ecol* 2002, **11**:197-214.
66. Hedrick P: **Large variance in reproductive success and the  $N_e/N$  ratio.** *Evolution* 2005, **59**:1596-1599.
67. Sambrook J, Russell D: *Molecular cloning: a laboratory manual* 3rd edition. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2001.
68. Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF: **M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods.** *Biotechniques* 2001, **31**:24-26.
69. De La Vega FM, Dailey D, Ziegler J, Williams J, Madden D, Gilbert DA: **New generation pharmacogenomic tools: a SNP linkage disequilibrium Map, validated SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies.** *Biotechniques* 2002:48-50. 52, 54
70. SAS: *SAS 9.1.3 Help and Documentation* Cary, NC: SAS Institute Inc; 2007.
71. **RECODE** [<http://watson.hgen.pitt.edu/register/docs/recode.html>]
72. Stephens M, Smith NJ, Donnelly P: **A new statistical method for haplotype reconstruction from population data.** *Am J Hum Genet* 2001, **68**:978-989.
73. Stephens M, Donnelly P: **A comparison of bayesian methods for haplotype reconstruction from population genotype data.** *Am J Hum Genet* 2003, **73**:1162-1169.
74. Weir BS: *Genetic Data Analysis II: Methods for Discrete Population Genetic Data* Sunderland, MA: Sinauer Associates Inc. Publishers; 1996.
75. Devlin B, Risch N: **A comparison of linkage disequilibrium measures for fine-scale mapping.** *Genomics* 1995, **29**:311-322.
76. Sved JA: **Linkage disequilibrium and homozygosity of chromosome segments in finite populations.** *Theor Popul Biol* 1971, **2**:125-141.
77. Hill WG: **Linkage disequilibrium among multiple neutral alleles produced by mutation in finite population.** *Theor Popul Biol* 1975, **8**:117-126.
78. Weir BS, Hill WG: **Effect of mating structure on variation in linkage disequilibrium.** *Genetics* 1980, **95**:477-488.
79. McVean GA: **A genealogical interpretation of linkage disequilibrium.** *Genetics* 2002, **162**:987-991.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

