BARD

FINAL REPORT

PROJECT NO. IS-2316-93

**Mapping the Chicken Genome Including Loci
Affecting Traits of Economic Importance**

M. Soller, H. Cheng, L. Crittenden

1998

Date: June 11, 1998

BARD Project Number: IS-2316-93C

Project Title: Mapping the chicken genome including loci affecting traits of
economic importance

Investigators:

Affiliated Institution:

Morris Soller

The Hebrew University of Jerusalem

Hans Cheng

Avian Disease and Oncology Laboratory

Lyman Crittenden

Avian Disease and Oncology Laboratory

Project Start Date: September 1994

Type of Report: Final

Signature

Principal Investigator

Signature

Authorizing Official

Dr. ONN ARRAD
DIRECTOR
AUTHORITY FOR RESEARCH
AND DEVELOPMENT

Table of Contents	Page
1. Abstract	2
2. Objectives of the original research proposal	3
3. Relevant Data, methodology, results and discussion	
3.1 Creating and enriching the chicken genome map	4
3.2 The full-sib intercross line	9
3.3 The Jerusalem Resource Population	15
3.4 Selective DNA pooling for QTL mapping	19
3.5 QTL mapping in the Jerusalem Resource Population	25
3.6 References	27
4. Description of the cooperation	30
5. Evaluation of research achievement with respect to the original objectives	30
6. Conclusions	31
7. List of publications	31
8. Tables and Figures	34 to 55

1. Abstract

A total of 195 microsatellites were added to the chicken genome map. Mapping of fifty known genes revealed a high degree of conserved linkage order between human and chicken genomes. A new, statistically powerful mapping design, the full-sib intercross line (produced by mating two parents, and intercrossing their progeny over a number of generations), was developed for use in species with high reproductive capacity. The Jerusalem Resource Population (JRP), now at the F12 generation, was established to implement this design in chickens. The biometrical picture in the JRP is similar to that generally found in chicken populations; inbreeding effects were not observed. The F2 and F3 generations of the JRP were genotyped with respect to twelve production traits, using a battery of 23 microsatellite markers. The number of significant effects was twice that expected on chance alone, validating the high statistical power of the JRP with respect to QTL differentiating the parental lines.

Selective DNA pooling, based on estimation of marker allele frequencies in pooled DNA samples, has been proposed to reduce high genotyping costs of QTL mapping. A method to correct for overlapping shadow bands of dinucleotide microsatellite markers in pooled DNA samples was developed and validated. In a retrospective study using this procedure, previously mapped loci affecting Marek's disease were successfully identified.

2. Objectives of the original research proposal

2.1 Develop a molecular genetic map of the poultry genome based on highly polymorphic DNA genetic markers.

2.2 Develop methods for saturating particular chromosomal regions with polymorphic genetic markers.

2.3 Continue the development of a resource population that can be used for fine-structure mapping of QTL affecting traits of economic importance in chickens.

2.4 Use the genetic map and resource population as the basis for identifying and gross mapping of chromosomal regions in the chicken genome that contain QTL affecting growth-rate and the associated traits of reproductive performance, body composition, and immune response.

2.5 Saturate a limited number of chromosomal regions containing QTL of major effect for traits of economic importance, as a step towards fine-structure mapping of QTL in these regions with a long-term goal of eventually identifying and cloning them by positional methods.

Objectives 2.1, 2.3, and 2.4 were actively pursued (see sections 3.1, 3.3, 3.5). Objective 2.2 was not pursued in view of rapid development of overall map saturation. Objective 2.5 was not pursued since the finding of major synteny between chicken and human genomes provides alternative routes to identification of mapped QTL. In place of these objectives we developed a general approach to QTL mapping in segregating populations (section 3.2) and a methodology for linkage mapping of QTL through selective DNA pooling with dinucleotide microsatellite markers (section 3.4).

3. Relevant Data, methodology, results and discussion

3.1 Creating and enriching the chicken genome map

The primary objective of the work on the US side of the BARD proposal was the generation of a detailed molecular genetic map of the chicken. The most convenient genetic material for building such a map, consists of an F2 or BC population derived from a cross between inbred lines, although other family based options are also workable. Fortunately, the availability of inbred lines and the large reproductive capability of the chicken make it an ideal subject, among agricultural animals, with which to generate such a "reference map" population. In the case of the East Lansing (EL) reference family, produced at ADOL, a single male from the inbred UCD-001 Jungle Fowl (JF) line was mated to a single female from the inbred UCD-003 White Leghorn (WL) line to produce F1 progeny (Crittenden et al., 1993). Two F1 males were individually backcrossed to 10 and 8 UCD-003 WL females, to produce 208 and 192 progeny, respectively. Large quantities of blood and DNA from each animal were stored for future use. Due to financial and technical limitations the mapping panel in common use consists of a subset of only 52 BC progeny derived from a single F1 male mated to four inbred WL females.

Another reference mapping family was generated in parallel by Nat Bumstead (Compton, United Kingdom). In the Compton (C) population, a single line-151 male was mated to a single line N female to produce progeny. Unlike the EL reference population, a single F1 female (heterogametic) individual was backcrossed to a line-151 male to generate the reference family (Bumstead and Palyga, 1992). The consequence of using a female (ZW) instead of a male (ZZ) F1 in the backcross is that the Z chromosome cannot be mapped in the C reference family. Conversely, the pseudoautosomal region of the W chromosome can be mapped in the C family but not in the EL family. DNA from a panel of 56 individuals forms the primary C mapping panel.

A third chicken reference family was generated more recently through a collaborative effort between Martien Goenen (U. of Wageningen, The Netherlands) and Euribrid, a poultry breeding company. Using two commercial broiler lines, 10

F2 families containing a total of 456 progeny were produced (Groenen et al., submitted). Although the Wageningen (W) population DNAs are not routinely available, the W population and resultant map have the advantage that virtually the whole F2 population has been mapped so that the genetic fine structure of the W map is superior to that available elsewhere. As a result, the W map is especially useful in bringing all the available data together when generating a consensus map (Groenen, Bumstead, and Cheng, in preparation).

Genetic markers are the other major component that is required to build a genetic map. Map development is additive when multiple groups genotype markers on a common reference family. Consequently, DNAs from the internationally-recognized reference families (EL and C) have been distributed to any interested group. Table 1 shows the major laboratories associated with either the EL or C maps. Each group genotypes their markers on the reference family and sends the data to the map curator (Hans Cheng for EL, and Nat Bumstead for C) for marker placement. The information is then made publicly available through the chicken genome database (Dave Burt, <http://www.ri.bbsrc.ac.uk/>).

Since the inception of the EL map in October 1991, the number of markers has grown dramatically each year as shown in Figure 1. The latest EL genetic map is shown in Figure 2.

A primary thrust of the ADOL under the BARD program has been to develop genetic markers for the EL map. In common with man and other species, two classes of marker, denoted Type I or Type II, have been developed. Type I markers are known genes (i.e., with a characterized function or phenotype) which provide "anchors" for building comparative maps, as genes are likely to be conserved across species. Type I genes may also provide candidates for QTL effects. Type II markers generally, but not always, exist in so-called "anonymous" DNA segments which do not encode a functional protein and therefore are under little or no evolutionary constraint against high levels of allelic diversity. These are primarily markers exhibiting polymorphisms that can be readily mapped and used in genome-wide QTL searches in a variety of relevant populations.

At ADOL, about 50 genes relevant to viral oncogenesis or with mammalian homologs were cloned, sequenced, and mapped using allele-specific oligonucleotides in PCR. Our results combined with those from others make it possible to carry out preliminary comparisons between the chicken and human genomes. Based on 154 genes, 36 human-chicken and 35 mouse-chicken conserved segments were defined. Using this sample of conserved segments, it is estimated that there have been 61 rearrangements since the common ancestor of chickens and human, which translates into 81 conserved segments between chicken and human. In contrast, there are 148 conserved segments between mouse and human, and 175 conserved segments between chicken and mouse. These results suggest that the avian genome has been relatively stable over time, almost 10-fold more than expected. From a comparative view, we reach the unexpected conclusion that the organization of the human genome is closer to that of the chicken than to that of the mouse!

This remarkable state of affairs has major implications for those trying to isolate genes encoding a QTL. The high degree of conserved linkage order indicates that by using the comparative map, uncharacterized chicken genes may be rapidly identified from the known locations of their homologs in the human genome. Thus, rather than having to clone the gene through chromosome walking, candidate genes may be revealed purely by their chromosomal position. If so, this combination of conserved gene synteny and order, along with the positional candidate gene approach, will save a great deal of time and effort. The fact that it is difficult, if not impossible, to resolve QTL to less than a 10 cM interval (roughly 10 million base pairs in mammalian genomes, about a third of that in the chicken genome), also indicates that mechanisms other than positional cloning alone will be required to isolate QTL-encoding genes.

Several classes of Type II markers have also been mapped. ADOL has focussed primarily on microsatellite markers. Microsatellites are tandem repeats of 1 to 6 bp motifs. For animal species, the ability to automate the genotyping along with the high likelihood of finding a DNA polymorphism have made microsatellite markers the primary genetic marker to map, and subsequently use in genome-wide

QTL searches. In this labor-intensive effort, using genomic and microsatellite-enriched libraries, 405 unique microsatellites have been identified which have resulted in 195 mapped microsatellites so far on the EL map.

The actual distribution of each type of marker across the EL map is shown in Table 2. As of May 1998, the EL map contained 954 markers, of which 926 markers described 42 linkage groups covering over 3,400 cM. There are 168 Type I genes mapped, most using RFLP, SSCP, physical linkage to a microsatellite, or allele-specific oligonucleotide primers with PCR. By far the most common Type II marker (446 entries) is the microsatellite. The other Type II markers are RFLP, minisatellite, AFLP, and RAPD. These are often found in genetic maps of other species. The lone exception is the CR1 marker, which utilizes PCR and primers to the chicken repeat element 1 to generate polymorphic products (Levin et al., 1994).

The chicken karyotype consists of 8 large chromosomes (macrochromosomes), 30 indistinguishable small chromosomes (microchromosomes), and the sex chromosomes, Z and W (Bitgood and Shoffner, 1990). The genetic map basically reflects the genomic architecture, as shown in Table 3; Ch. 1 is slightly smaller than predicted by its DNA content (14.8% vs. 20.5% of the genome). With regard to possible distribution bias for a given marker type, microsatellites are very infrequent in the Z chromosome (relative to expectation based on chromosome size alone), and moderately infrequent in the microchromosomes. Minisatellites are clustered, and are much more frequent in the microchromosomes. This may reflect the preferential location for these elements near the centromeres and telomeres, which presumably constitute a greater percentage of microchromosomal DNA. For reasons unknown, AFLP and RAPD markers are less frequent in the microchromosomes. While there are not many CR1 markers, there is an obvious preference for the macrochromosomes. In examining the chromosomal distribution of these Type II markers, one needs to keep in mind the selection process that goes into their identification. For example, usable microsatellites must be surrounded by single copy DNA and thus will be rare in clusters of repetitive DNA. To a lesser extent, usable CR1 and RAPD primer sites cannot be identified in highly repetitive DNA, since a useless smear of PCR products is then produced.

Gene density as indicated by Type I markers, does not show an obvious chromosomal preference when measured by genetic distance. However, many uncertainties remain in comparing genetic length to physical distance across the genome, and the ratio between the two is likely to be quite variable. For example, Rodionov et al. (1992) estimate that the microchromosomes exhibit ~3 times more recombination per cytogenetic length than do macrochromosomes. The data in Table 3 are superficially not in agreement with this hypothesis, as probable microchromosome linkage groups constitute about 31.8% of the map length in cM vs. 27.5% of the physical map from cytogenetic measurements. However, some of the smaller linkage groups may not actually be on microchromosomes and some microchromosomes may not yet be identified by two or more markers. Both of these factors underrepresent the proportion of the genetic map associated with the microchromosomes. In addition, most telomere regions are probably not represented in the genetic map (due to lack of a nearby marker), and these missing regions probably constitute a higher percentage of the true microchromosomal genetic map. Furthermore, a few genotyping errors probably remain in the data, which would artificially inflate recombination proportions, and the cytogenetic distance may also contain some systematic bias. In summary, until clone-based physical maps are available, most questions of non-random gene and marker density will remain uncertain. The available data in Table 3 do not rule out the hypothesis of McQueen et al. (1996) that gene density (per kb of DNA) is higher in microchromosomes than in macrochromosomes.

Many of the ADOL microsatellites have been utilized in other chicken populations. Primarily because of the microsatellite markers scored in the EL, C, and W reference families, the genetic maps could be aligned and compared. The EL and W maps do not show a major systematic difference, but the corresponding C linkage groups are consistently larger. This is unlikely to be due to the different sex in which the C and EL meioses occurred. Recombination in the two sexes can be directly compared in the W map, and, as in other species, the genetic map of the heterogametic sex (female) is shorter than the genetic map of the homogametic sex (male) but only slightly so (1.1%).

these species, this poses a major block to linkage mapping and positional cloning of QTL, and to the implementation of marker-assisted selection.

A full-sib intercross line (FSIL) is produced by mating two parents, and intercrossing their progeny over two or more generations to form a large "mapping population". The genetic makeup of the mapping population is thus completely determined by the chromosome constitution of the two founder parents, and tends, therefore, to a high degree of linkage disequilibrium between marker alleles and QTL alleles. The FSIL design thus provides a QTL mapping population, for utilization in outcrossing species, with power and other attributes similar to those obtaining for populations derived from a cross between inbred lines.

In the context of the BARD program we designed and analyzed the statistical power of a FSIL design, according to the specific marker and QTL composition of the founding parental pair, and compared this to the power of F₂, full-sib and half-sib designs. Potential applications of FSIL designs for positional cloning and marker-assisted selection were also explored.

Results

Assuming absence of dominance at the QTL, the statistical expressions for the power of the various experimental designs for determining marker/QTL linkage, can all be expressed in the form $Nd^2\delta^2$; where N is total population size, d is the standardized QTL gene effect, and δ^2 , denoted the "design parameter", is a function of the experimental design and of marker and QTL composition of the parents of the mapping population. This means that within a particular design, all combinations of population size and gene effect which yield the same value of Nd^2 , will have the same power. This enables the power of various designs to be compared on the basis of the relative Nd^2 values required for equivalent power.

In order to provide a convenient basis of comparison of the various designs, Nd^2 values for power from 0.10 to 0.90, are provided for the inbred line design at Type I error, $\alpha=.01$ (Table 4). Relative Nd^2 values, denoted RN values, are given for all other designs, as a multiple of the inbred line Nd^2 values necessary to obtain equivalent power. As an example of the use of Nd^2 values, we note that from the

given Nd^2 value for power of 0.90, it follows that this power will be obtained at $N=100$, $d=0.59$; at $N=500$, $d=.27$; and at $N=1000$, $d=.19$.

We now consider the relative Nd^2 (RN) required with the FSIL design to provide the same power as the inbred line design, according to marker and QTL composition of the founder parental pair. Taken over all informative marker and QTL compositions, most RN values are in the range 1.0 to 4.0, at any given power, with a modal value of 2.2. Thus, it seems reasonable to conclude that on the average a FSIL design will deliver power equivalent to that of an inbred line design, at twice the corresponding population size or Nd^2 .

However, this "average" view does not exhaust the possibilities of the FSIL design. The reason for this, is that some of the marker combinations can be transformed into other more powerful marker combinations, according to the specific marker-QTL coupling relationships in the parental chromosomes. For example, denoting marker alleles, M1, M2, and QTL alleles A1 and A2, marker/QTL parental composition $M1A1/M1A1 \times M1A2/M2A1$ (denoted $M_{1112}A_{1121}$) has an RN of 11.64 for power of 0.90; while marker/QTL combination $M_{1112}A_{1112}$ has an RN of 1.33. Yet by judicious search, a marker having distribution $M_{1121}A_{1121}$ can be found, which will provide an RN of 1.33. Thus, by judicious choice of markers or by combining marker alleles having similar associated effects, it should be possible to transform any parental pair having an A_{1112} QTL composition into an $M_{1112}A_{1112}$ combination, with an RN of 1.33; and any parental pair having an A_{1122} QTL composition, into an $M_{1122}A_{1122}$ combination, with an RN of 1.00. In the analysis of an FSIL design, therefore, one would initially use a fairly high Type I error, to identify potential marker-QTL linkages; and then subject these to a second round of marker search and analysis, attempting to identify a marker combination that markedly increased statistical significance. In this way, it should be possible to obtain from a FSIL design population, power virtually equivalent to that of an inbred line design.

For the half-sib and full-sib family designs, as more families are tested at given total population size, the individual family is necessarily smaller. This tends to reduce statistical power. For the half-sib design, the effect is not great, and RN values remain in the range of 4 to 10, with a modal value of about 6. For the full-sib

design, RN values for given power increase strongly with the number of families tested. Thus, RN values for 25 families for a power of 0.50, range from 4.8 to 7.9, depending on marker composition; while those for 200 families at the same power, range from 12.1 to 19.4. Thus, effective RN for a full-sib family design may be closer to 10 overall; depending however, on marker composition and number of families tested.

Discussion

The results of this analysis show that in principle the FSIL design brings to QTL mapping in outbreeding populations (animals, fruit trees, some plants), many of the advantages that the inbred line designs provide for selfing species, or species for which inbred lines are available. Foremost among these advantages is the greater statistical power for given population size and gene effect, relative to the alternative full-sib or half-sib designs applicable in outbreeding populations. A useful attribute of the FSIL design, is the ability to achieve major increase in power, in the same sample and data base, by searching for more informative marker combinations in chromosomal regions showing indications of QTL presence. This is not relevant for the half-sib design, and impractical and theoretically complex for the full-sib design. Following marker adjustment, a FSIL design can consistently provide six to ten fold greater power than equivalent half-sib or full-sib family designs. In addition, FSIL mapping populations, being in strong population-wide linkage disequilibrium, are amenable to mapping through selective DNA pooling (Darvasi & Soller, 1994; Lipkin *et al.*, 1997). This can provide major reductions in genotyping costs of the initial genome scan for QTL of interest.

Yet another advantage of the FSIL design, is the ability to continue the population for a number of generations allowing phenotyping and genotyping data to be accumulated over the generations, a matter which can be of practical importance in experimental situations, where rearing facilities are limited. Continuing a FSIL population over a number of generations by repeated intercrossing, however, will increase the cumulative number of recombination events between closely linked sites. Consequently, a FSIL design requires a more dense marker spacing

deleterious effects of the resultant inbreeding, might best be circumvented by initiating a number of independent F1 families from independent parental pairs drawn from the source population; intermating each of these F1 families, and continuing with the F2 or F3 families that shows least signs of inbreeding depression. Aside from the inescapable inbreeding in the F1 generation, inbreeding in later generations is much reduced simply by maintaining a large effective number of parents, and can be further minimized by deliberate choice of parental mating pairs to avoid inbreeding.

A FSIL design can also be viewed as a procedure for general identification and mapping of QTL, not necessarily related to a specific commercial breeding population. In this case, the information on QTL effects and map location obtained from the FSIL design can feed into an MAS program in a commercial population in a number of ways: (i) By focussing QTL mapping efforts in the commercial population on chromosomal regions that are likely to contain QTL. (ii) By serving as a step to linkage disequilibrium QTL mapping in the commercial population. Specifically, because of the mapping precision possible with a FSIL design, it is possible to saturate the region containing the QTL with markers, aiming at a spacing of 1 to 2 cM (so that on the average a QTL would be only .25 to .50 cM from the nearest marker), and then look for marker associated effects on a population-wide basis (Lande & Thompson, 1990). (iii) By serving as a source of "marked" QTL for introduction into the commercial line by marker-assisted introgression. In particular, as in an inbred line design, a FSIL design provides information on specific marker-QTL phase relationships. Consequently, once a FSIL design has located a QTL of interest to a specific chromosomal region, flanking and spanning markers can be defined, so that favorable QTL alleles are defined within specific haplotypes of marker alleles. Such marker-QTL complexes could be treated as quasi-Mendelian super-alleles, and serve as a source of favorable QTL alleles for marker-assisted introgression into the commercial population.

In addition to the above, a FSIL population, can be created as an adjunct to a commercial breeding population, using founder-animals from the breeding population itself. In this case, mapping information from the FSIL design population

would be directly applicable to the commercial population. This could be used as a basis for linkage disequilibrium mapping or marker-assisted introgression, as above. Multi-generational BLUP based methods for identification of marker-QTL phase, when QTL effect and location are known are under active development (Fernando & Grossman, 1989; Meuwissen & Goddard. 1996). Thus, one can envision a program in which a FSIL design population is established, while simultaneously DNA samples and data are accumulated in the commercial breeding population. Mapping information obtained in the FSIL design population, would provide a basis for retroactive genotyping of the commercial population leading to BLUP estimates of marker-QTL phase in the chromosomes of the commercial population.

3.3 The Jerusalem Resource Population (HU)

The Jerusalem Resource Population, was established to implement these propositions in practice, and to generate a population, based on FSIL principles, for use in mapping QTL of economic importance in chickens. In order to provide a wide genetic base, the population was based on a cross between a broiler line male and females of a layer line. The broiler male was taken from a normal commercial breeding population, and hence represented a single individual chosen from a population that should be basically in linkage equilibrium. Instead of a single layer line female, however, the second parent was represented by a set of five females taken from a semi-inbred layer line (Line-0). This was done in order to produce a large F1 population, for rapid expansion of the F2. Later studies showed that about half of the chromosomal regions in Line-0 are represented by only one haplotype; others are represented by anywhere from two to four haplotypes (as shown by the number of different marker alleles found among the five female founder individuals). The effective degree of inbreeding was therefore about 50%. On the average, therefore, the five founder females represent the equivalent of a single female taken from a random mating population. Thus, the effective situation among the founder parents of the JRP should approximate that proposed for a chromosomal-design population.

The first generation individuals were hatched in 1989, and the population has been reproduced successfully by continued intercrossing, over the course of eleven generations, being now (June 1998) at the F12 generation. The present study, however, is limited to analysis of data through the F9 generation. In each generation, ten to 30 males, and 35 to 85 females were taken to reproduce the population as shown in Table 5. Taking the harmonic mean, through the F9 generations, the effective number of males was 12.2, and effective number of females was 57.0, so that inbreeding will have accumulated at a rate of $\Delta F = 1/8N_m + 1/8N_f = 0.012$ per generation (Falconer, 1996). Thus, total inbreeding in the population will be $F = 0.25$, for the first generation of brother x sister mating; $F_t = 1 - (1 - \Delta F)^t = .092$, for the remaining eight generations (t). This is a relatively small increment. Total inbreeding, will then be $F = 0.25 + (.092)(.75) = 0.32$.

The population did not show any obvious effects of inbreeding. Hatchability, although not specifically recorded in each generation, was sufficient to reproduce the population on the basis of single hatch following a three week egg collection period. There were no obvious deformities or recessive traits that came to expression in the chicks. The measured phenotypic traits were more or less stable over the course of the nine generations, and did not show any trends which could be attributed to inbreeding depression. Thus, these results show the feasibility of establishing and maintaining a FSIL population.

The data and sample base accumulated for the JRP through the F9 generation, illustrate very nicely the ability of a FSIL population to accumulate data over the course of generations. Thus, although only a total of 200 to 250 individuals (males plus females) were hatched in any one year; over the eight years of the study data were accumulated on many hundreds of individuals. This was true even for the body composition traits recorded on males only, and the egg production traits recorded on females only.

When a population is divided into a collection of full-sib families, biometrical principles show that half of the genetic variance in the population will be found within the families, and half will be found between the families. Thus, when a FSIL population is based on a single parental pair drawn from the same random mating

population, one would expect it to comprise half of the genetic variation originally present. In the case of the JRP, the male and female parents were drawn from parental lines that differed widely in many of their phenotypic traits, and that also represented historically distinct breeds. Thus, part of the genetic variance lost by sampling only a single individual (the broiler line male) or utilizing a semi-inbred line (the founder females), should have been retrieved by the fact that the JRP was initiated by sampling from phenotypically and historically distinct populations. This variance component is analogous to the genetic variance generated in the F₂ of a cross between two inbred lines. Thus, we would expect genetic parameters of the JRP to resemble those found within the classical breeds of chicken.

Because there was so little further inbreeding following the establishment of the F₂ generation, it was expected that there would not be any noticeable changes in phenotypic and genetic parameters of the JRP across generations. Indeed, as noted above distinct time trends were not observed for the descriptive parameters of any of the recorded traits. The same holds true for phenotypic correlations between the traits, and for parent-offspring correlations. Thus, it is possible to use the pooled descriptive statistics and parent-offspring correlations over all eight recorded generations, as indicative of overall population biometry.

When this is done, the biometrical picture obtained from the JRP is basically similar to that generally found within broiler or layer populations. Coefficient of variation for body weight was on the high side, but plausible considering that this was the trait for which the parental lines differed most greatly. Coefficient of variation for abdominal fat pad weight or proportion of abdominal fat pad weight were very high; those for relative breast weight and shank length were much less. These results are consistent with those reported in the literature. Coefficient of variation of age and weight at first egg were moderate, and similar to those reported in the literature.

The coefficient of variation of post-peak percent lay was very high, much greater than generally reported for percent lay. This trait also showed marked skewness in the direction of lower values. The high coefficient of variation is probably a function of the skewness. Both of these parameters suggest that percent

lay in the experimental conditions of this experiment, was subject to uncontrolled environmental factors or non-additive that affected it greatly.

The phenotypic correlations between the various traits confirm the impression of normal trait behavior. There were strong positive correlations between juvenile body weight, breast weight, abdominal fat pad weight, shank length, and weight at first egg; positive correlations between juvenile body weight and proportion of breast weight or abdominal fat pad weight, but a negative correlation with relative shank length. All of these are as expected from the part-whole relationships of body weight and its components, or relative growth patterns of these components and correspond to results commonly obtained with these traits. There was a positive correlation between age and body weight at first egg, as expected and commonly obtained.

Heritability estimates for immune response, juvenile body weight, and for age and body weight at first egg correspond to those commonly reported in the literature, and attest to significant genetic variation for these traits in the JRP. Although the structure of data did not allow calculation of parent-offspring correlations for the body composition traits, the parent-offspring correlations of these traits with juvenile growth rate were all highly significant. This is an indirect confirmation of a genetic basis for the observed variation in these traits. Thus, the heritability estimates and the parent-offspring correlations between traits are consistent with the presence of considerable genetic variation in the traits of interest, within the existing database of the JRP. The exception to this appears to be the very low heritability of percent lay, which we attribute to the high coefficient of variation and strong negative skewness of this trait. The reasonable estimates obtained for heritability and genetic correlations, also support the general validity of the pedigree recording procedures in this population.

For 22 markers allele frequencies in the F3 generation, in two-thirds of the cases, corresponded reasonably closely (within ± 0.15) with those in the parental individuals. None of the alleles present in the original parental individuals were lost from the population; in only one instance, was an allele originally present at a moderately high frequency (>0.20), brought to a frequency below 0.10; while in a

number of instances alleles originally present at low frequencies, increased in frequency. Thus, the general impression is that allele frequency changes are due to random sampling, rather than to linkage to lethal or sublethal alleles, although the latter possibility cannot be eliminated.

When Mendelian transmission of marker alleles from parent to offspring was examined, nine markers examined showed less than 10% of false genotypes. The Mendelian behavior of these markers supports the accuracy of genotyping for these markers. Beyond this, the Mendelian behavior of these markers also supports the accuracy of the pedigree records, at least for the F2 and F3 generations.

Thus, on the basis of the biometrical and marker analyses, we conclude that the JRP represents a genetically representative chicken population, within which QTL affecting immune response and a variety of body composition traits of economic importance are segregating, and which is not subject to severe inbreeding depression or other effects associated with inbreeding. With respect to egg production traits, there is good evidence for segregation of QTL affecting age and body weight at onset of sexual maturity; less evidence for segregation of QTL affecting percent lay, as measured in the conditions of this population. The general reasonable behavior of the data base, also support the validity of the procedures used for data and pedigree recording in this population. Thus, the existing data and sample base of the JRP, and future data and samples to be obtained from it, should be able to provide a basis for QTL mapping.

3.4. Selective DNA pooling for QTL mapping (HU and ADOL)

Previous theoretical work in our laboratory and elsewhere had shown that most of the information on marker-QTL linkage was concentrated in the high and low phenotypic tails of the mapping population (Darvasi and Soller, 1992). Later theoretical studies in our laboratory showed that virtually all of the mapping information was found in the difference in marker allele frequencies between the high and low phenotypic tails (Darvasi and Soller, 1994). This provides an opportunity to carry out QTL mapping by estimation of marker allele frequencies in pooled DNA samples taken from the individuals making up the high and low phenotypic tails. Through this procedure, termed "selective DNA pooling" the high costs of screen-

ing large populations for marker allele frequencies can be greatly reduced (Darvasi and Soller, 1994; Lipkin et al., 1998). At present, most of the markers used in constructing genomic maps are dinucleotide microsatellites having a poly(TG) motif (Cheng et al., 1995). These are typed by PCR. Alleles, different in repeat number, are resolved as discrete bands by electrophoresis through an acrylamid sequencing gel. Estimating allele frequencies in pooled DNA samples is based on a linear relation between the initial number of copies of the allele in the pool, as determined by the allele frequency in the group of individuals making up the pool; and the final allele band intensity, as determined by densitometry. However, this estimate is often confounded by the presence of "shadow" or "stutter" band, i.e., artifactual PCR products, derived from the genomic poly(TG) tract by deletion or insertion of one or more repeat motifs (Hauge and Litt, 1993). Consequently, an observed band in the pool can be a composite of a main product of the genomic tract and a number of shadow products generated by contiguous alleles in the same lane. In our work under the BARD program, we first studied the relation of shadow band intensity to repeat number of the genomic tract from which it was derived, and of the number of deleted or inserted motifs by which it differs from the genomic tract. We then constructed a model in which this linear function was used to correct the densitometric intensity of microsatellite bands for overlapping shadow bands. The model was validated in cooperation with ADOL.

Relative intensity of shadow bands

Statistical methods. Given the densitometric values of main and shadow bands, observed relative intensity, RI, of a given shadow band for a given allele, was calculated as:

$$RI_{n,i} = D_{n,i}/D_n$$

where,

n is the number of repeats in the native genomic tract of the allele A_n ;

i is the "order" of the shadow band, i.e., the number of inserted ($1=+1$) or deleted ($1=-3$ to -1) motifs by which it was derived from the genomic tract;

— $RI_{n,i}$ is the relative intensity of the i^{th} shadow band derived from the genomic tract of A_n ;

D_n is the densitometric intensity of the main band derived from the genomic tract of A_n ;

$D_{n,i}$ is the densitometric intensity of the i th shadow band derived from the genomic tract of A_n .

Densitometry measurements for the seven microsatellites were obtained for individual samples that were homozygous for a given allele; or for which the two alleles were well separated. RI were calculated for the first three leading ($i=-3$ to -1), and the first trailing ($i=+1$) shadow bands. Regression equations for RI on n for the various shadow band orders, were as follows:

$$RI_{n,-1} = -.184 + .046 n$$

$$RI_{n,-2} = -.323 + .037n$$

$$RI_{n,-3} = -.212 + .025n$$

$$RI_{n,+1} = -.129 + .020n$$

The total amount of shadow product was remarkable: for $n=24$, total shadow product was twice that of the main band of the allele.

Allele frequency estimates in pooled DNA samples

Shadow correction. Given the D_n values of all bands in a pool, and taking into account shadow bands, the following model was used to estimate the shadow-corrected intensity (CI) of all bands:

$$CI_n = D_n - \sum (CI_{n-i} RI_{n-i}),$$

where,

n , D_n and i are as above, and i runs from -3 to $+1$;

CI_{n-i} is the shadow-corrected intensity of D_{n-i} ;

$RI_{n-i,i}$ is the RI of that shadow band of A_{n-i} that overlaps the main band of A_n .

This is the shadow band that is derived by deletion or insertion of i repeats from A_{n-i} .

For example, if $n=20$ and $i=-2$, then $RI_{n-i,i}$ would be the RI of the shadow band derived by a deletion of two repeats from the genomic A_{22} allele.

Thus, $CI_{n-i} RI_{n-i,i}$ is the estimated intensity of the i^{th} shadow band derived from the allele A_{n-i} ;

For a pool with k bands there will be k equations with k unknown CI's. These can readily be solved as a set of simultaneous equations. Correcting for the above four shadow bands was sufficient to quantify microsatellite PCR products.

Total allele products. Since the shadow band products are derived by further amplification of the products of the genomic tract, they also represent the quantitative products of the alleles present in the pool. Consequently, the total product derived from the genomic tract consists of the product in the main band, plus the product in the derived shadow bands. Because RI increases with n , a higher proportion of the total product derived from long as compared to short alleles, is found in the shadow bands as compared to the main band. Consequently, CI values for the main band taken alone tend to underestimate the frequency of long alleles relative to short alleles. Hence, the sum of the intensities of the shadow bands, estimated from the above RI equations, was added to the CI_n , giving:

$$CT_n = CI_n \sum RI_{n,i}$$

where,

i runs from -3 to $+1$;

CT_n is the sum of all allele products;

$RI_{n,0} = 1.0$, so that $CI_n RI_{n,0}$ is the shadow-corrected intensity of the A_n main band.

Allele frequency estimates in the pools. Given CT values for all bands in the pool, the shadow corrected frequency estimate of allele A_n in the pool was calculated as:

$$F_n = CT_n / \sum CT_n$$

where,

n runs from 1 to k

$\sum CT_n$ is the sum of CT values of all k bands in the lane.

Pool densitometry and Selective DNA pooling

Two experiments were performed to validate the ability of densitometry to correctly estimate allele frequencies in pooled samples. These involved pools of purified DNA constructed to have specific allele frequencies; and pools of fresh or

frozen-defrosted red blood cells representing a specific line of chicken. A third experimental was carried to retrospectively validate the ability of pooled DNA samples to uncover marker/QTL linkage.

Experimental

Constructed samples. Purified DNA was extracted from individual blood samples, and genotyped with respect to markers ADL020 and MCW5. Individuals were then chosen, with respect to each markers, to make up a series of pools spanning a range of allele frequencies. Nine pools were constructed with respect to marker ADL020, and three pools with respect to marker MCW5. Each pool was prepared once.

Red blood cell pools. Fresh red blood samples were taken from a single line of chicken, and individually genotyped with respect to markers ADL115, LeiO75, LeiO88, Lei163. The individual blood samples were then frozen, defrosted, and pooled again. DNA was extracted from the pooled samples. Each pool (fresh or frozen) was prepared twice. The pools were then densitometrically genotyped for the markers.

Selective DNA pooling. Two DNA pools were prepared from frozen-defrosted individual blood samples of the 40 most resistant and 40 most susceptible birds, out of 272 F2 individuals used previously at ADOL to map QTL affecting resistance to Marek's disease. Each pool was prepared in two independent replicates; each replicate preparation was amplified twice, to a total of four amplifications per pool. Means of the four amplifications of each pool were compared to estimates based on individual genotyping. In addition, the differences in allele frequency between resistant/susceptible pools were used in a retrospective test to identify markers showing linkage to QTL affecting Marek's susceptibility. The pools were genotyped for 15 markers chosen from among those used in the original mapping study. For each marker, one individual from each of the parental lines was genotyped and used to estimate the allele repeat number. Shadow correction of the pool data was based on the repeat number of the reported sequence when available, or the genotypes of the above individuals. Linkage to a QTL was determined by the difference in parental line allele frequencies in the two pools. The test

of significance was based on Khatib et al. (1994) and Lipkin et al. (1998).

Results

Constructed DNA pools. The regressions of the pool estimates on the individual frequencies are shown in Table 6 and Figure 3. The results were similar for both markers. Densitometric estimates based on raw densitometry had very high correlations with the individual frequencies, but were strongly biased, with overestimation of rare alleles and underestimation of frequent alleles (Figure 3a). The intercept and the regression coefficients were significantly different from 0.0 and 1.0, respectively. Shadow-corrected estimates, on the other hand, were evenly distributed about the individual frequencies, with the intercept and the regression coefficients not different from 0.0 and 1.0, respectively (Figure 3b). All correlation coefficients were higher for the corrected estimates than for the raw estimates. In addition, the uncorrected estimates included a large number of "false" positive frequency estimates, indicating presence of rare alleles in the pools which were not present on individual genotyping. Such errors were almost completely absent in the shadow-corrected data.

Red cell pools. Results are shown in Table 6 and Figure 4. For the first three markers, raw- and shadow-corrected densitometric estimates showed a high correlation with estimates based on individual genotyping; with the corrected densitometric estimates having higher correlation and less bias than the raw densitometric estimates. For the marker Lei0163, the shadow-corrected estimates, though unbiased, had a relatively low correlation with the individual estimates (Table 6). For this marker, the peaks of the heterozygous individuals indicated differential amplification of the long allele, such that the relative intensity of the main bands of the two alleles was 0.38, instead of the expected 1.00. Applying this coefficient to the shadow corrected value improved the accord between the pools and the individual frequencies.

For all markers, regression and correlation coefficients obtained with pools constructed from fresh or frozen-defrosted red blood cells, were substantially the same.

Retrospective study of selective DNA pooling. For twelve of the markers of this study, the correlation between allele frequency estimates based on shadow-corrected pool densitometry and those based on individual genotyping was 0.96; estimates were unbiased (intercept and slope of regression line not differing from 0.0 and 1.0, respectively). Three of the markers (MCW0082, ADL0345, ADL0161) showed less accord with individual genotyping, or gave biased results even after shadow-correction (Table 7). The reasons for this are still under study.

Generally there was good agreement between the difference (D) in parental allele frequencies between high and low pools, as estimated by pool densitometry, and the difference based on individual genotyping (Figure 3, $r=0.85$, and regression coefficients not differing from 0.0 and 1.0).

Pooled over all markers, the error component due to replicate preparations of the same pool (σ^2_{TR}) was equal to $= 0.00003$; the error component due to replicate amplifications of the same pool-replicate (σ^2_{TA}) was equal to 0.00029 . Taken together with the binomial component of the variance (σ^2_B), this gave a standard error, SE(D), for the difference between two tails equal to 0.0064 . On the basis of this, tests of significance of differences between pools were performed. Figure 5 shows comparative significance test results based on pool densitometry and on individual genotyping for three chromosomes. Both pool genotyping and individual genotyping indicate the presence of QTL affecting resistance to Marek's disease near LEI0101 on chromosome 1, and between ADL0171 and MCW0305 on chromosome 8.

3.5 QTL mapping in the Jerusalem Resource Population (HU and ADOL)

All individuals of the F2 and F3 generations of the JRP were genotyped with respect to a battery of 23 microsatellite markers covering a major portion of the chicken genome. Chromosome 1 markers were also genotyped in the F4 individuals. About half of the markers were genotyped at HU and half at ADOL. For purposes of analysis, the progeny were designated as carrying zero, one or two alleles of paternal origin, or as having two alleles of indeterminate origin (where the allele was found in both the founder sire and one or more of the founder dams

of the JRP). The analysis was performed using the GLM ANOVA procedure of SAS, with sex, generation, and marker genotype as fixed effects. A total of 23 markers and 12 traits were analyzed. The traits, and number of trait measurements included in the analysis are given in Table 8. The significance level of marker/trait associations are given in Table 9.

Out of a matrix having a total of 23 (markers) x 12 (traits) = 276 tests of significance, 11 P values fell in the range $.105 > P > .055$; 23 in the range $.055 > P > .015$, and 9 in the range $.015 > P$; expected numbers in each class, based on Type I error alone, are 13.8, 11.0, and 4.14, respectively. Thus, the number of P values in the range $.105 > P > .055$ was no greater than expected by chance alone. Hence, these must represent primarily sampling variation, rather than true marker/trait associations. Thus, associations at this significance level, will not be considered further.

In contrast, the number of P values in the range $.055 > P$ was more than twice that expected on chance alone (32 as compared to 14). Hence, for the most part, these must represent true marker/trait associations, and are taken to be indicative of marker/QTL linkage.

The number of significant effects per markers ranged from zero (7 markers) to 5 (one marker). The distribution of significant effects among markers was more or less as expected on chance alone. However, the six chromosome 1 markers as a group had significantly fewer significant effects per marker (a total of 5, or .83 per marker) than the remaining 17 markers (a total of 27 significant effects, or 1.58 per marker). This corresponds with other reports that the very large Chromosome I has a paucity of functional genes, relative to the other chromosomes.

The number of significant effects per trait varied from 0 (2 traits) to 6 (2 traits). In this case, the distribution of effects differed significantly from that expected on chance alone. Examination of the traits showed a modest tendency for the number of significant effects to follow heritability. Thus, abdominal fat pad weight and shank length, both having high heritability, had six significant effects each, while percent γ -lay, with very low heritability, had only a single significant effect. However, there were many exceptions: mature body weight, with high heritability had zero signifi-

cant effects, while immune response with low heritability had four. At this stage, one cannot attach too much meaning to these different patterns. They may be due to lack of power or uneven distribution of markers.

Taken as a whole, these results validate the basic concept of the Jerusalem Resource Population, and show its high statistical power with respect to QTL differentiating the parental lines. Additional statistical analyses, are required, to determine the actual effects associated with the various alleles, and their distribution among paternal and maternal chromosomes. Data and DNA samples are available for eleven generations of this population. Extension of the present study to include additional generations and additional markers can be expected to uncover many more QTL, and provide more accurate map locations for them. Taken together with the extensive human/chicken synteny, the way will then be open to identification of candidate genes responsible for the observed effects.

3.6 References

- Beckmann, J.S. & Soller, M. (1988). Detection of linkage between marker loci and loci affecting quantitative traits in crosses between segregating populations. *Theor. Appl. Genetics* **76**: 228-236.
- Bitgood, J.J. & Shoffner, R.N. (1990). Cytology and Cytogenetics. In: Poultry Breeding and Genetics (ed. R.D. Crawford), pp. 401-427. Elsevier, New York.
- Bumstead, N. & Palyga, J. (1992). A preliminary linkage map of the chicken genome. *Genomics* **13**: 690-697.
- Cheng, H.H., Levin, I., Vallejo, R.L., Crittenden, L.B., Hillel, J., Didgson, J.B., and Soller, M. 1995. Development of a genetic map of the chicken with high-utility markers. *Poultry Science* **74**: 1855-1874.
- Crittenden, L.B., Provencher, L., Santangelo, L., Levin, I., Abplanalp, H., Briles, R.W., Briles, W.E. & Dodgson, J.B. (1993). Characterization of a Red Jungle Fowl by White Leghorn backcross reference population for molecular mapping of the chicken genome. *Poultry Science* **72**: 334-338.

- Darvasi, A. (1997). Interval specific congenic strains (ISCS): An experimental design for mapping a QTL into a one centimorgan interval. *Mammalian Genome* **8**: 163-167.
- Darvasi, A. & Soller, M. (1992). Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theor. Appl. Genet.* **85**: 353-359.
- Darvasi, A. & Soller, M. (1994). Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics* **138**: 1365-1373.
- Darvasi, A. & Soller, M. (1995). Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* **141**: 1199-1207.
- Falconer, D.S. & Mackay, T.F.C. (1997). Introduction to Quantitative Genetics, 3rd Edition. Longman Group Ltd. Essex, England.
- Fernando, R.L. & Grossman, M. (1989). Marker-assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.* **21**: 467-477.
- Haley, C.S., Knott, S.A. & Elsen, J.M. (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**: 1195-1207.
- Hauge, X.Y. & Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Human Molecular Genetics* **2**: 411-415.
- Khatib, H., Plotsky, Y. & Soller, M. (1994). Estimating relative microsatellite allele frequencies in pooled DNA samples. *PCR Methods Applications* **4**: 13-18.
- Knott, S.A., Elsen, J.M. & Haley, C.S. (1996). Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theor. Appl. Genet.* **93**: 71-80.
- Knott, S.A. & Haley, C.S. (1992). Maximum-likelihood mapping of quantitative trait loci using full-sib families. *Genetics* **132**: 1211-1222.
- Lande, R. & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**: 743-756.
- Lander, E.S. & Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185-199.

- Levin, I., Santangelo, L., Cheng, H., Crittenden, L.B., & Dodgson, J.B. 1994. An autosomal genetic linkage map of the chicken. *J. Heredity* **85**: 79-85.
- Lipkin, E., Mosig, M.O., Darvasi, A., Ezra, E., Shalom, A., Friedmann, A. & Soller, M. (1998). Mapping loci controlling milk protein percentage in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers. *Genetics* (in press).
- McQueen, H.A., Fantes, J., Cross, S.H., Clark, V.H., Archibald, A.L. & Bird, A.P. (1996). CpG islands of chicken are concentrated on microchromosomes. *Nature Genetics* **12**: 321-324.
- Meuwissen, T.H.E. & Goddard, M.E. (1996). The use of marker haplotypes in animal breeding schemes. *Genet. Sel. Evol.* **28**: 161-176.
- Rodionov, A.V., Myakoshina, Yu.A., Chelysheva, L.A. et al. (1992). Chiasmata on the lampbrush chromosomes of *Gallus gallus domesticus*: cytogenetic study of recombination frequency and length of linkage groups. *Genetika (Moscow)* **25**: 53-63.
- Soller, M. (1994). Marker-assisted selection, an overview. *Animal Biotechnology* **5**: 193-208.
- Soller, M., Brody, T. & Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Applied Genet.* **47**: 35-39.
- Soller, M. & Genizi, A. (1978). The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. *Biometrics* **34**: 47-55.
- Van der Beek, S.J., van Arendonk, J.A.M. & Groen, A.F. (1995). Power of two- and three-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families. *Theor. Appl. Genet.* **91**: 1115-1124.
- Weller, J.I., Kashi, Y. & Soller, M. (1990). Daughter and granddaughter designs for mapping of quantitative trait loci in dairy cattle. *J. Dairy Science* **73**: 2525-2537.

4. Description of the cooperation

In addition to joint planning and evaluation of experiments, the two laboratories cooperated actively in mapping the F2 and F3 generations of the Jerusalem Resource Population (JRP). ADOL provided microsatellite markers that were used at HU, and in addition ADOL did the actual genotyping of about ten markers on the HU samples. Samples were amplified at HU and then sent by express mail to ADOL for genotyping, using automatic sequencer machine at ADOL.

The two laboratories also cooperated actively in the study of selective DNA pooling. The HU laboratory developed methodology for correcting densitometric estimates of band intensity by shadow band overlap. The two laboratories then designed a joint experiment using samples prepared at ADOL to validate this methodology. Dr. Ehud Lipkin of HU, travelled to ADOL for two weeks to work on this together with ADOL staff, and the two laboratories are jointly analyzing the data.

5. Evaluation of research achievement with respect to the original objectives

Much of the original objectives were achieved. The chicken genome map was greatly enriched in the course of the three years of the proposal (section 3.1). A theoretical basis for the JRP (The full-sib intercross line design) was fully explored (Section 3.2), and the validity of the JRP itself as a mapping population was established by biometrical analyses and by marker analyses (Section 3.4). (Section 3.5). Numerous putative QTL affecting a variety of traits of economic importance were identified (Section 3.5). These will provide a firm basis for further QTL mapping in the JRP.

In addition to the original objectives, synteny of chicken and human genome was explored and found to be remarkably conserved (Section 3.1), and a methodology for implementing selective DNA pooling using dinucleotide microsatellite markers was developed and validated (Section 3.4).

6. Conclusions

1. There is remarkably conservation of genome organization between chicken and human. This provides a firm basis for comparative positional cloning of QTL affecting economic traits in chicken.
2. The Jerusalem Resource Population is an effective resource for mapping chicken QTL affecting traits of economic importance.
3. Selective DNA pooling, with correction for shadow bands, is a useable methodology for QTL mapping. This can allow QTL mapping to be extended to half-sib family designs in chicken and other species, without incurring excessive genotyping costs.
4. A full-sib intercross family design, promises greatly increased QTL mapping power for animal species with high reproductive capacity such as chicken, swine or fish.

7. List of publications

- Khatib H, Genislav E, Crittenden LB, Bumstead N, Soller M (1993). Sequence tagged microsatellite sites as markers in chicken reference and resource populations. *Animal Genetics* 24:355-362.
- Darvasi A, Weinreb A, Minke V, Weller JI and Soller M (1993). Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* 134: 943-951.
- Khatib H, Soller M (1993). Mononucleotide repeat polymorphism at the GGY locus in chicken. *Animal Genetics* 24: 397-98.
- Darvasi A, Soller M (1994). Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. *Theoretical and Applied Genetics* 89:351-357 .
- Khatib H, Plotsky Y, Soller M (1994). Estimating relative microsatellite allele frequencies in pooled DNA samples. *PCR Methods and Applications* 4:13-18.

Khatib H, Soller M (1994) . Dinucleotide repeat polymorphism at the TENS and SCHOX1 loci in chicken. *Animal Genetics* 25:199.

Soller M, Marker-assisted selection, an overview (1994). *Animal Biotechnology* 5:193-208.

Darvasi A, Soller M (1994). Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics* 138: 1365-1373.

Shalom A, Darvasi A, Barendse W, Cheng H, Soller M (1996) . Single parent cosegregant pools for allocation of markers to a specified chromosomal region in outcrossing species. *Animal Genetics* 27:9-17.

Song JZ, Soller M, Genizi A. (1998). The full-sib intercross line (FSIL) design: A QTL mapping design for outcrossing species. *Genetical Research* (in press).

Soller M, Andersson L. (1998). Genomic approaches to improvement of disease resistance in farm animals. In: *Genetic resistance to animal diseases*. OIE Scientific and Technical Review, Volume 17 (1) April 1998.

Cheng, H.H. and Crittenden, L.B. 1993. Microsatellite markers for the chicken genetic map. *Poultry Sci.* 73: 539-546.

Cheng, H.H., Levin, I., Vallejo, R.L., Crittenden, L.B., Hillel, J., Dodgson, J.B., and Soller, M. 1995. Development of a genetic map of the chicken with high-utility markers. *Poultry Science* 74: 1855-1874.

Levin, I., Santangelo, L., Cheng, H., Crittenden, L.B., & Dodgson, J.B. 1994. An autosomal genetic linkage map of the chicken. *J. Heredity* 85: 79-85.

Smith, E.J., Cheng, H. and Vallejo, R. (1996). Mapping functional chicken genes: an alternative approach. *Poultry Sci.* 75: 642-647.

Dodgson, J.B., Cheng, H.H. and Okamoto, R. (1997). DNA-marker technology: a revolution in animal genetics. *Poultry Sci.* 76: 1108-1114.

Cheng, H.H. (1997). Mapping the chicken genome. *Poultry Science* 75: 1101-1107.

Smith, E.J. and Cheng, H.H. (1998). Mapping chicken genes using preferential amplification of specific alleles. Micro. and Comp. Genom. 3: 13-20.

Cheng, H.H. and Dodgson, J.B. (1988). Chicken genomics. in: Animal Biotechnology: Molecular Breeding and Transgenics. International Academic Publishers, in press.

Table 1. Laboratories associated with the chicken genetic map

Bernard Benkel	Centre for Food and Animal Research Ottawa, Ontario, Canada
Nat Bumstead	Institute for Animal Health Compton, England
Terry Burke	University of Leicester Leicester, England
Dave Burt	Roslin Institute Midlothian, Scotland
Hans Cheng	Avian Disease and Oncology Laboratory East Lansing, MI, USA
Lyman Crittenden	Michigan State University East Lansing, MI, USA
Jerry Dodgson	Michigan State University East Lansing, MI, USA
Martien Groenen	Wageningen Agricultural University Wageningen, The Netherlands
Jossi Hillel	The Hebrew University of Jerusalem Rehovot, Israel
Susan Lamont	Iowa State University Ames, IA, USA
Abel Ponce de León	University of Massachusetts Amherst, MA, USA
Marcia Miller	Beckman Research Institute of the City of Hope Duarte, CA, USA
Eugene Smith	Avian Disease and Oncology Laboratory East Lansing, MI, USA
Morris Soller	The Hebrew University of Jerusalem Jerusalem, Israel
Hideaki Takahashi	National Institute of Agrobiological Resources Ibaraki, Japan
Alain Vignal	Centre INRA de Toulouse Castenet Tolosan, France

Table 2: East Lansing Genetic Map (5/98)

Linkage Group	No. of Markers	cM	Classical	Genes ¹	Micro	Mini	RFLP	AFLP	RAPD	CRI	Other
1	185	498		24 (5)	87	2	11	40	6	14	1
2	112	434		13 (6)	57	1	5	24	7	7	1
3	92	328	1	14 (3)	35		5	19	10	6	2
4	70	213		10 (2)	33		1	14	5	5	2
5	43	145		8 (1)	16		2	9	2	2	1
6	32	111		6 (4)	10		1	4	3	2	
7	29	128		4 (1)	19			12		3	
8	33	94		5 (2)				9			
16	11	50	1	5		4			1		
7	59	242	1	11 (1)	15	1	1	10	16	4	1
W	9	13	1				2	5			
F04	5	14	2	1							
E16	12	43		1	3			6	1	1	
E17	3	21					1		2		
E18	12	28		3 (1)	7				2		
E21 (Ch. 77)	4	12		2	1		1				
E22	6	64	1	2 (1)	1	2					
E24	2	0		1							
E25	5	29		2		1					
E26	7	46			4	1					
E27	6	11	1			7					
E29	21	45		3	9		3	2	1	2	1
F30	17	47			13			1	2	1	
F31	8	14		4 (1)	4						
F32	3	2		2			1				
E34 (Ch. 57)	5	12		1	1			2	1		
F35	8	43		1	5		2				
F36	22	90		2	14		1	5			
F38	5	20	1	3 (1)	1						
F40	2	11				2					
F41	10	59		5	5						
F46	12	82		1 (1)	4	4				2	1
F47	6	13			6			6			
F48	22	79		4 (2)	8			2	2		2
F49	9	54	1	2	3						
F50	4	38			1	2					
F52	5	32		2	1		1				
F53	10	83		2	6			1			1
F54	6	71		5							
F57	2	14		1		1					
F58	2	8				1					
F59	6	35		3 (3)	2	1					
Unlinked	26			5 (3)	12	3	1	2	3		
42	954	3,421	10	158 (38)	408	36	39	176	67	47	14

¹ The number of genes linked to microsatellites is given in the ().

Table 3: Distribution and Spacing of Markers on East Lansing Genetic Map (5/98)											
Linkage Group	No. of Markers	cM ¹	% Genome ²	Micro ^{3,4}	Mini	RFLP	AFLP ⁴	RAPD ⁴	CR1 ⁴	Genes ^{4,5}	Other
1	185	498 (14.8%)	20.5	92 (5.4)	2	11	40 (12.5)	6 (83.0)	14 (35.6)	24 (20.8)	1
2	118	434 (12.9%)	12.8	63 (6.8)	1	5	24 (18.1)	10 (43.4)	7 (62.0)	13 (33.4)	1
3	92	328 (9.7%)	9	38 (8.6)		5	19 (17.3)	10 (32.8)	6 (54.7)	15 (21.9)	2
4	70	213 (6.3%)	7.1	35 (6.1)		1	14 (15.2)	5 (42.6)	5 (42.6)	10 (21.3)	2
5	43	145 (4.3%)	5.8	22 (6.6)		2	9 (16.1)	2 (72.5)	(-)	8 (18.1)	1
6	32	111 (3.3%)	3.2	20 (5.6)		1	4 (27.8)	3 (37.0)	2 (55.5)	6 (18.5)	
7	29	128 (3.8%)	3.2	11 (11.6)			12 (10.7)	(-)	3 (42.7)	4 (32.0)	
8	33	94 (2.8%)	1.9	21 (4.5)			9 (10.4)	(-)	(-)	5 (18.8)	
16	11	50			4		(-)	1 (50.0)	(-)	6 (8.3)	
Z	59	242 (7.2%)	7.1	16 (15.1)	1	1	10 (24.2)	16 (-)	4 (60.5)	12 (20.2)	1
W	9	13 (0.4%)	1.9	1 (13.0)		2	5 (2.6)	(-)	(-)	1 (13.0)	
E04 - E59 ⁶	264	1071 (31.3%)	27.5	113 (9.7)	25	9	26 (42.0)	10 (107.1)	6 (182.7)	56 (19.6)	6
Unlinked	26			12	3	1	2	3		5	
42	954	3,376	100	446 (7.6)	36 (93.8)	39 (86.6)	176 (19.2)	67 (50.4)	47 (71.8)	168 (20.1)	14

¹ The number in the () is the percentage of the total genetic coverage.

² Based on weights of chromosomes taken from cutouts of photomicrographs of mitotic cells.

³ Does not include microsatellites linked to genes.

⁴ The number in the () is the average genetic distance per marker type.

⁵ Includes both classical genes and DNA markers.

⁶ Assumes that these are microchromosomes but does not include E21 and E34.

Boxes that are shadowed indicates values outside the norm taking certain biases into consideration.

Table 4. Inbred line design. Power as a function of Nd^2 . N, population size; d, gene effect at the QTL

Nd^2	Power
4.63	.10
10.00	.27
16.48	.50
20.00	.61
35.20	.90
40.00	.94
90.00	1.00

Table 6. Regression and correlation of estimates of allele frequencies in the pools, obtained by pool densitometry, on estimates obtained from individual genotyping of all members of the pool, according to marker, substrate, and correction procedure¹.

Exp.	Marker	n	Sub.	Proc.	Intercept	Slope	Correlation
1	ADL020	36	DNA	R	0.026**	0.895**	0.998
				C	0.003	0.988	0.998
	MCW512	DNA	R	0.105**	0.579**	0.958	
				C	0.000	0.998	0.994
	All	50	DNA	R	0.044**	0.824**	0.981
				C	0.002	0.991	0.997
2	ADL115	17	FB	R	0.022	0.911	0.986
				C	-0.005	1.022	0.999
			FDB	R	0.019	0.922	0.988
				C	-0.003	1.010	1.000
	Lei075	16	FB	R	0.054	0.784*	0.967
				C	0.007	0.971	0.997
			FDB	R	0.050	0.800	0.946
				C	-0.007	1.028	0.997
	Lei088	28	FB	R	0.071**	0.503**	0.928
				C	0.013	0.912*	0.986
			FDB	R	0.068**	0.523**	0.924
				C	0.009	0.939	0.991
	Lei163	16	FB	R	0.088*	0.648*	0.940
				C	0.034	0.862	0.796

			CD	0.003	0.989	0.995
		FDB	R	0.086*	0.655*	0.932
			C	0.038	0.848	0.764
			CD	0.005	0.981	0.995
All	77	FB	CD	0.005	0.978	0.995
		FDB	CD	0.002	0.992	0.996

*, $p < 0.05$; **, $p < 0.01$

¹Exp. 1, pools of purified DNA, constructed to cover a wide range of allele frequencies;

Exp. 2, pools of fresh or frozen-defrosted blood samples representative of a single stain of chicken.

n, number of data points.

Sub., substrate for the pooling: DNA, purified DNA; FB, fresh blood; FDB, frozen-defrosted blood.

Proc., Correction procedure: R, raw data; C, corrected for shadow bands; CD, corrected for shadow bands and differential amplification (see text for details).

All, data pooled over all markers, separately by experiments.

Table 7. Selective DNA pooling. Regression and correlation of pools frequency estimates, on estimates obtained from individual genotyping.

Chr.	Marker	n	Intercept	Slope	Correlation
1	LEI0101	8	0.016	0.936	0.995
1	UMA1.220	8	-0.005	1.022	0.986
1	ADL0268	8	0.064	0.743	0.901
1	LEI0108	8	-0.001	1.003	0.987
2	MCW0082	8	0.022	0.912	0.752
2	ADL0152	6	-0.008	1.023	0.943
2	ADL0185	6	0.067	0.799	0.942
2	LEI0086	8	0.018	0.927	0.914
8	ADL0345	6	0.347**	-0.042**	0.245
8	ADL0161	4	0.640**	-0.280**	0.983
8	ADL0171	6	0.003	0.991	0.991
8	ADL0258	4	0.027	0.947	0.990
8	MCW0305	8	0.004	0.986	0.965
8	ADL0322	8	0.014	0.945	0.945
All		102	0.037*	0.875*	0.903

*, $p < 0.05$; **, $p < 0.01$

n, number of data points.

All, data pooled over all markers.

Table 8. Traits, and number of data points per trait.

Trait acronym	Trait description	Data points		
		F2	F3	F4
IR1	Immune response	156	NA	NA
JW7	7-wk body weight	158	188	162
JW13	13-wk body weight	158	188	162
JWS	Carcass weight	51	55	49
BW	Breast weight	51	55	49
AW	Abdominal fat pad weight	51	55	49
SL	Shank length	51	55	49
AL	Age at first egg	61	84	60
WL	Body weight at first egg	61	84	59
EW	Weight of first egg	55	84	59
MW	Mature body weight	61	74	68
PL	Percent Lay	61	74	68

Table 9A. Tests of significance (P values) of marker/trait associations for immune response, juvenile body weight and male body composition in the Jerusalem Resource Population. IR, immune response; JW7, body weight at 7 weeks of age; JW13, body weight at 13 weeks of age; JWS, body weight at slaughter; BW, breast weight; AW, abdominal fat pad weight; SL, shank length.

Marker	IR1	JW7	JW13	JWS	BW	AW	SL
MCW5	.04	N	N	N	N	.04	N
ASMA	N	N	N	N	N	N	N
MCW4	N	N	N	N	N	N	N
TENS	N	.002	.004	N	N	N	N
M154	N	N	N	N	N	N	N
GGY	N	.07	N	N	N	N	.03
B100	N	N	N	N	N	N	N
B122	N	.07	.03	.01	.01	.02	.08
M3131	N	N	N	N	N	N	N
M41	N	N	N	N	N	N	N
M21	N	N	N	.06	N	.02	.07
M602	N	N	N	N	N	N	.02
CHOX	N	N	N	N	N	N	N
M301	N	.04	N	N	N	N	N
CMYC	.05	.003	N	.07	N	.05	.01
M151	.01	N	N	N	N	N	.02
VITC	.04	N	N	N	N	.05	N
B55	N	.02	N	N	N	N	N
B266	.07	N	N	N	N	N	N
ADL19	N	N	N	.07	N	.005	N
B149	N	N	N	N	N	N	N
N721	N	N	N	N	N	N	N
H722	N	N	N	N	N	N	.05

Table 9B. Tests of significance (P values) of marker/trait associations for female reproductive performance and mature body weight, in the Jerusalem Resource Population. AL, age at first egg; WL, body weight at first egg; EW weight of first egg; MW, mature female body weight, PL, percent lay post peak.

Marker	AL	WL	EW	MW	PL
MCW5	N	N	.04	N	.06
ASMA	N	N	N	N	N
MCW4	N	N	N	N	N
TENS	N	N	N	N	N
ADL154	N	N	N	N	N
GGY	N	N	N	N	N
B100	.05	N	N	N	N
B122	.04	N	N	N	N
M3131	N	N	N	N	N
M41	N	N	N	N	.10
M21	N	N	N	N	N
M602	N	N	N	N	N
CHOX	N	N	N	N	N
M301	N	N	N	N	N
CMYC	N	N	N	N	N
M157	.04	N	.01	N	N
VITC	N	N	.03	.06	N
B55	N	N	N	N	N
B266	N	N	N	N	N
ADL19	N	N	N	N	.02
B149	N	.07	N	N	N
H721	N	N	N	N	N
H722	.04	N	N	N	N

Figure 1: Progression of East Lansing Genetic Map

Ch. 7

* Includes microsatellite markers only

Figure 3. Scatter diagram and regression lines of estimates of allele frequencies obtained by pool densitometry (y), on estimates obtained from individual genotyping of the members of the pool (x). —, expected regression line if pool estimates are an unbiased estimate of the frequency based on individual genotyping; ---, regression line of pools on individual estimates; ■, estimates based on pools.

- A. Experiment 1, constructed pools using purified DNA samples. A1. Raw data, $y=0.04+0.82x$, $r=0.981$. A2. Shadow-corrected data, $y=0.00+0.99x$, $r=0.997$.
- B. Experiment 2, pools using fresh blood and frozen-defrosted blood. B1. Raw data, $y=0.05+0.74x$, $r=0.995$. B2. Shadow-corrected data, $y=0.00+0.98x$, $r=0.996$.

Figure 4. Results of the significance tests for the selective DNA pooling study. P values (to the $-\log_{10}$) of difference in parental line allele frequency between high and low pools, as estimated by pool densitometry or individual genotyping, for markers on Chromosomes 1, 2, and 8. —, pool genotyping; -----, individual genotyping.

Figure 3A1. Constructed pools, raw data

Figure 3A2. Constructed pools, shadow-corrected data.

Figure 3B1. Pools of fresh or frozen-defrosted blood, raw data

Figure 3B2. Pools of fresh or frozen-defrosted blood, shadow-corrected data

Figure 4

Figure 4

Figure 4