Chapter 2
Statistical and Methodological Considerations in Exercise Genomics

Heather Gordish-Dressman and Joseph M. Devaney

Keywords Additive model • Analysis of covariance • Analysis of variance • Base pair • Candidate single nucleotide polymorphism • Chi square • Codominant model • Confounding • Dominant model • Epistasis • Expression quantitative loci • Genome-wide association studies • Genomic control • Genetic risk score • Hardy–Weinberg equilibrium • Lewontin's D' • Linkage disequilibrium • Logistic regression • Multiple testing • Odds ratio • Population stratification • Power calculations • Correlation coefficient (r^2) • Recessive model • Sample size calculations • Single nucleotide polymorphism • Structured association • Transcription factor 7-like 2 • Thyrotropin-releasing hormone receptor • Tag SNP • X chromosome • Y chromosome

Introduction

This chapter discusses the basic statistical concepts necessary for the analysis of genetic data. It focuses on single nucleotide polymorphism (SNP) data, a very active research area, and describes the methods used for data analysis and some important issues that arise specifically with genetic data. Information is given for studies using single SNPs, i.e., candidate gene studies, and for studies using multiple SNPs, such as genome-wide association studies (GWAS). Further information on GWAS describes how they can be a useful tool for academicians, clinicians, health/fitness professionals, and exercise scientists by detailing both what has been done and what is expected for the future. This chapter will provide the reader with an outline of things to consider when planning a study using genetic data and things to consider when reading a published report using genetic data.

H. Gordish-Dressman (✉)
Children's National Medical Center, Research Center for Genetic Medicine,
111 Michigan Avenue, N.W., Washington, DC 20010, USA
e-mail: hgordish@cnmcresearch.org

L.S. Pescatello and S.M. Roth (eds.), *Exercise Genomics*,
Molecular and Translational Medicine, DOI 10.1007/978-1-60761-355-8_2,

Statistical Considerations in Exercise Genomics

The statistical analysis of genetic data requires special considerations and encounters challenges not usually seen with other data types. These considerations include: calculation of Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD), additional requirements when performing power and/or sample size calculations, choosing the correct genetic model to analyze data, issues with nonautosomal loci (i.e., those on the X and Y chromosomes), and issues of confounding in the sample due to genetic factors. The statistical considerations in exercise genomics section of this chapter will focus on statistical techniques used to study genetic information in nonfamilial populations, specifically the analysis of single nucleotide polymorphism (SNP) data. Other methods to examine genetics in family groups (i.e., familial aggregation, segregation, and linkage analysis) will not be covered here, but are discussed in Chap. 1 of this book.

Analysis of nonfamilial populations falls into two major types, candidate gene association studies and their extension to GWAS. The candidate gene approach examines a limited number of associations with gene loci with known related functions that could be relevant to the phenotype of interest. GWAS, on the other hand, explore nearly all loci known to have a polymorphism to determine if any are associated with the phenotype of interest. The candidate gene approach is useful if there is prior knowledge of a gene, its function, and relationship to your phenotype. In contrast, GWAS are used when there is no or minimal prior knowledge and there is general interest in testing up to one million loci to see if any are associated with the phenotype. Both types of approaches use similar statistical methods, but GWAS in particular require additional considerations.

Hardy–Weinberg Equilibrium

The HWE principle states allele and genotype frequencies are in equilibrium (i.e., they remain constant from generation to generation unless other processes act to change them). The HWE formula predicts the expected genotype frequencies given their allele frequencies and statistically tests them against those observed in the sample. HWE assumes a large population size, random mating, and an absence of mutation, migration, or natural selection. In most populations HWE is assumed, as there is often no information on migration, mutation, or selection, and mating is random is an inherent assumption. When a biallelic locus shows deviation from HWE, it can mean that one or more of the HWE assumptions have been violated, but it more commonly suggests genotyping errors [1]. It is for this reason that each genetic locus is tested for HWE. The genotyping for any loci not in HWE should be verified. In addition, association tests using loci not in HWE can lead to biased estimates of the association's magnitude [2], yet another reason for testing each SNP for HWE.

The HWE formula ($1 = p^2 + 2pq + q^2$) is used to calculate expected genotype frequencies and a one degree of freedom Chi square (χ^2) test compares them to

the observed frequencies [3]. For a biallelic locus with alleles A and B, p^2 corresponds to the frequency of the homozygous AA genotype, q^2 to the BB genotype, and $2pq$ to the heterozygous AB genotype. Allele frequencies are calculated from these genotype frequencies, and finally the expected genotype frequencies are determined. A one degree of freedom χ^2 test compares the expected to observed genotype frequencies. The χ^2 tests the null hypothesis of no statistically significant difference in the distribution of genotype frequencies between the observed and expected. Therefore, a nonsignificant p value indicates HWE. A significant p value indicates a deviation from HWE, and the genotypes should be verified. HWE is one of the few statistical procedures where a nonsignificant p value is desirable.

Calculation and testing of HWE requires several steps. Given a SNP with the following genotype frequencies TT = 418, TG = 253, GG = 33, and total sample size = 704, the first step is to calculate allele frequencies from the observed genotype frequencies.

$$\begin{aligned}\text{Allele frequency of T}: p(\text{T}) &= ((\text{TT} * 2) + \text{TG}) / (\text{Total} * 2)\\ &= ((418 * 2) + 253) / (704 * 2)\\ &= 0.773,\end{aligned}$$

$$\begin{aligned}\text{Allele frequency of G}: p(\text{G}) &= ((\text{GG} * 2) + \text{TG}) / (\text{Total} * 2)\\ &= ((33 * 2) + 253) / (704 * 2)\\ &= 0.227.\end{aligned}$$

Second, the allele frequencies are used to calculate the expected genotype frequencies.

$$\text{Genotype frequency of TT} = p(\text{T})^2 * \text{Total} = (0.773)^2 * 704 = 421,$$

$$\begin{aligned}\text{Genotype frequency of GT} &= 2 * p(\text{T}) * p(\text{G}) * \text{Total}\\ &= 2 * 0.773 * 0.227 * 704 = 247,\end{aligned}$$

$$\text{Genotype frequency of GG} = p(\text{G})^2 * \text{Total} = (0.227)^2 * 704 = 36.$$

Finally, the observed genotype frequencies (418, 253, and 33) are compared to the expected frequencies (421, 247, and 36) using a 1 degree of freedom χ^2 test to yield $\chi^2 = 0.455$; $p = 0.4999$. Thus, in this case, the $p > 0.05$ indicates the SNP is in HWE, the desirable outcome.

Linkage Disequilibrium

LD is defined as the nonrandom association of SNPs located near each other in the genome [4] arising because certain sections of the genome tend to be inherited together more often than expected by chance. The closer together two SNPs are, the

less likely that a recombination event occurs between them. Therefore, the more likely they are inherited together. SNPs in high LD are almost always inherited together and can serve as proxies for one another [5, 6].

There are two major reasons why LD is important in any analysis of genetic data. The first relates to inference of causation. A strong association between a SNP and phenotype may indicate that particular SNP is the causal locus; alternatively, the true causal locus may be another SNP in high LD with the one tested. Two SNPs in high LD would be expected to show similar associations with a phenotype. Consequently, a SNP at one locus may appear to be associated with a phenotype, but only because it is in high LD with the true causal locus [7].

The second reason is that LD relates to the error rates of statistical tests, particularly their effect on the simultaneous analysis of many SNPs as in GWAS. All statistical tests have an error rate α, which is the rate at which the true null hypothesis of no difference is incorrectly rejected. In other words, the results of the test indicate there is a statistically significant difference when there is not. This alpha level (α) is conventionally set to 0.05, corresponding to a 5% chance of incorrectly rejecting the null hypothesis and stating a significant difference. For a single statistical test or even a handful of tests, this is a reasonable level of error most scientists and clinicians are comfortable with reporting. But when multiple statistical tests are performed on the same data (i.e., when testing associations among many SNPs and a phenotype), this error rate is no longer acceptable. For example, when performing 100 statistical tests with the same $\alpha = 0.05$, five rejections of the null hypothesis are expected to be solely due to chance. This becomes an even greater problem with GWAS, which attempt to test a phenotype against all known SNPs in the genome. Technology now allows us to simultaneously test one million SNPs (and that number will increase as technology moves forward). With the same $\alpha = 0.05$, 50,000 significant results due to chance alone would be expected. So from the resulting list of significant p values, 50,000 are false positives and are due to chance rather than a true association between the SNP and phenotype. Of course which 50,000 associations are false positives and which are true associations are not known, making interpretation of our results extremely difficult if not impossible. One solution to the problem of multiple SNP and phenotype testing is to perform fewer statistical tests (i.e., relying on LD).

Two SNPs in high LD are essentially giving the same information and are expected to show the same associations with the phenotype. There is no reason to test both SNPs against the phenotype. The total number of SNPs tested can be decreased by excluding redundant SNPs that are in high LD with each other. Most products available to perform GWAS have already taken LD between SNPs into consideration and have removed many redundant SNPs [8]. When performing multiple statistical tests on many SNPs from a candidate gene, the overall number of test being performed can be decreased by measuring LD between the SNPs beforehand. Testing for LD leads to a decrease in the overall error rate and an increased confidence in drawing the correct conclusions from the data.

There are several different measures of LD, each having its own characteristics. Those most commonly encountered are the r^2 (correlation coefficient) and D' (Lewontin's D'). For an excellent definition and comparison of LD measurements, see Devlin and Risch [9]. The calculations of both r^2 and D' use only allele frequencies

of the two SNPs and are related. If the frequencies of the alleles at two loci are defined as p_A, p_a, p_B, p_b, then, $D = p_{AB} - p_A p_B$. Because the sign of D depends on the arbitrary assignment of which allele is A and which is B, D' is more commonly used. $D' = |D/D_{max}|$ where D_{max} is the lesser of $p_A p_b$ or $p_a p_B$ if D is positive and the lesser of $p_A p_B$ or $p_a p_b$ if D is negative. A further common LD measure, r^2, is defined as $r^2 = D^2 / p_A(1 - p_A)p_B(1 - p_B)$ [7]. There are several resources available to calculate LD and a helpful list is given at http://www.genes.org.uk/software/LD-software.shtml (see Appendix).

D' has been deemed the most appropriate for fine mapping [9] because it is directly related to recombination rate, and it is the only LD measure not sensitive to allele frequencies. However, D' can be upwardly biased in small samples, especially those with rare alleles, since its magnitude is highly dependent on sample size. The term D' is also difficult to interpret when its values deviate from the extremes of 0 and 1 [10]. A value of $D' = 1$ indicates complete LD and a $D' = 0$ indicates no LD, but interpretation of intermediate values is difficult. Therefore, r^2 has been shown to have several properties making it more useful than D' [7]. The measure r^2 has more reliable sample properties at low allele frequencies (a common occurrence with SNPs), has the strongest relationship with population genetic theory, and has a simple linear relationship with sample size [7, 11]. r^2 is one of the most common LD measures reported, but exhibits sensitivity to the variation in allele frequencies across loci making the comparison of r^2 values between different locations (and different studies) problematic. As with D', an $r^2 = 1$ indicates perfect LD where the observations at one locus provide complete information about the other locus and $r^2 = 0$ indicates no LD. However, intermediate values are more intuitive since the value of r^2 is related to the amount of information provided by one locus about the other [10], just as the r^2 value from a linear regression provides the amount of variability of one variable described by the other. For example, if two SNPs have an LD of $r^2 = 0.43$, by knowing an individual's genotype for one of the SNPs, we then can confidently predict the genotype of the other SNP for only 43 out of the 100 individuals. Higher LD, and thus higher r^2 values, allows us to accurately predict the second SNP for more individuals.

The magnitude of LD measurements considered strong and weak has evolved largely by convention. Many define strong LD as an r^2 value larger than or equal to 0.80 [12]. Others define strong LD with confidence intervals where the one-sided upper 95% confidence interval bound on D' is greater than or equal to 0.98 [3].

Power and Sample Size Considerations with Genetic Data

Making estimates of sample sizes and/or power is an important part of any proposed experiment. Not only are they usually required by funding agencies as part of grant application, they are also important for the researcher to ensure adequate testing of the hypothesis of the sample of interest. The methods used to calculate sample size and power estimates are the same with genetic data as with other traditional data types, but allele frequencies need to be taken into consideration when using genetic data. Typical sample size calculations yield a number of subjects, usually per group,

that are necessary to detect a significant difference between those groups. If a calculation indicates a sample size of 50 per genotype group is required, the allele frequency of the minor allele then needs to be considered to calculate how many total subjects you need to get 50 heterozygous mutants (the rarest genotype). For example, a sample size calculation might determine that you need 50 homozygous wild-type individuals (AA) and 50 homozygous mutant individuals (BB) to adequately test a hypothesis. If the SNP has a major allele frequency of 0.85 [p(A)] and a minor allele frequency of 0.15 [p(B)], it would require a total sample size of 2,500 to ensure there are 50 homozygous mutants in the sample. One can determine the number of individuals of each genotype directly from the allele frequencies according to HWE, just as one does for HWE calculations. Given 100 individuals, the number of those of each genotype is given by:

$$\#\mathrm{AA} = p(\mathrm{A})^2 * \mathrm{total} = 0.85^2 * 100 = 72.25 = 72,$$

$$\#\mathrm{AB} = 2 * p(\mathrm{A}) * p(\mathrm{B}) * \mathrm{total} = 2 * 0.85 * 0.15 * 100 = 25.5 = 26,$$

$$\#\mathrm{BB} = p(\mathrm{B})^2 * 100 = 0.15^2 * 100 = 2.25 = 2.$$

So with a minor allele frequency of 0.15, only two individuals out of every 100 would have a homozygous mutant genotype. Therefore, 2,500 subjects are needed to ensure that there are 50 homozygous mutants in the sample. The allele frequency of the minor allele greatly increases the total number of subjects needed for the study and is an additional important consideration.

Choosing the Most Appropriate Genetic Model

For any biallelic locus (with alleles A and B), there are three possible genotypes (AA, AB, and BB), and each allele can differentially influence the phenotype. In other words, the A allele can dominate the B or vice versa. This penetrance of each allele leads to different genetic models. The first model is the codominant model where all three genotypes have a different effect on the phenotype, and if those correspond to a dose–response with the effect of the phenotype increasing (or decreasing) with the number of mutant alleles, it is considered an additive model. Sometimes the heterozygote genotype's effect on the phenotype is more similar to either the homozygous wild-type or the homozygous mutant genotypes. In this case if the heterozygotes and homozygous wild-type individuals show similar outcomes, a recessive model can be used where the homozygous mutants (BB) are tested against the combination of the heterozygotes and homozygous wild types (AA + AB). If the heterozygotes are similar to the homozygous mutants, a dominant model can be used (AA vs. AB + BB).

Examples of three SNPs, each tested against a phenotype where each SNP demonstrates a different genetic model, are now described. In the first instance, the mean

relative difference in one repetition maximum (1RM) dynamic muscle strength was compared to a genetic variant in the Resistin gene (i.e., *RETN* −420 C>G) where C is the major allele and G is the minor allele. The relative differences in 1RM means ± SEM for each genotype were CC = 37.2 ± 3.0%, CG = 31.7 ± 2.7%, and GG = 21.3 ± 5.3% [12]. Each of the means appears to be independent, and there seems to be a dose–response with a decreased mean difference with the G allele. In this case, the data show a codominant (additive) model and are best analyzed as three genotypes. A comparison of subcutaneous fat volume with the insulin-induced gene 2 (i.e., *INSIG2* − 101,025 G>C) (G = major allele, C = minor allele) showed the following mean ± SEM: GG = 243,473 ± 5,713 mm^3, GC = 269,331 ± 5,645 mm^3, and CC = 263,941 ± 10,940 mm^3 [13]. In this case, the means for the GC and CC genotypes are much more alike than the homozygous wild-type GG genotype, indicating that this locus may follow a dominant model. Combining the AC and CC genotype together and comparing the AA genotype, a dominant model would be a good choice for analysis. Lastly, a comparison of isometric strength with rs3739287 (A = major allele, G = minor allele) showed the following: AA = 63.6 ± 1.2 kg, AG = 62.7 ± 2.8 kg, and GG = 103.7 ± 5.5 kg (unpublished data). Here the AA and AG genotypes are very similar and both very different from the GG genotype, indicating that this locus may follow a recessive model. Analyzing as a recessive model, combining the AA and AG and comparing to GG would be the best choice here.

Recessive models are the most difficult to test from a sample size perspective, i.e., if the SNP has a low minor allele frequency, the total number of homozygous mutants (BB) may be too low to effectively test. The codominant and/or additive models have the additional need to adjust resulting *p* value for multiple comparisons because we are comparing three genotype groups.

X and Y Chromosome SNPs

Most genetic variations for association studies are located on one of the 22 autosomal chromosomes, but those on the X and Y chromosomes have special considerations [14]. For SNPs on autosomal chromosomes, each individual has two copies of the allele, one inherited from each parent. Therefore, genotypes report two alleles for each subject (AA, AB, or BB). SNPs located on the X chromosome are similar in females having two X chromosomes, i.e., each female has two alleles for each X chromosome loci. Males, on the other hand, are termed hemizygous, having only one X chromosome and only one allele. Depending on the platform performing the genotyping, X chromosome loci in males are often represented in a similar fashion to autosomal loci and show two alleles, while in reality, they only have one. So a male with an A allele at a X chromosome locus will be represented by a genotype of AA, and a B allele represented as BB even though there is truly only one allele. Normal males cannot be heterozygous for an X chromosome locus. Therefore, a heterozygote genotype usually indicates a genotyping error, or more rarely, a misclassification of gender or a genetic condition such as Kleinfelter's syndrome where the man has more than one X chromosome. Similarly, females

cannot have genotypes for Y chromosome loci. Calculation of HWE for a nonautosomal chromosome locus is usually done in females only for an X chromosome locus and in males only for those on the Y chromosome locus.

Analysis of Multiple SNPs – Epistasis

Often, especially when studying complex traits or diseases, we have several SNPs that together may affect our phenotype. One SNP may mask or alter the effect of another. Epistasis refers to the phenomenon where the effects of one gene are modified by one or several other genes [15]. The classic example of epistasis comes from Bateson in 1909 [15]. In the following table, we see the phenotype (hair color) in mice with known genotypes at two loci, 1 and 2.

	Genotype at locus 2		
Genotype at locus 1	g/g	g/G	G/G
b/b	White	Gray	Gray
b/B	Black	Gray	Gray
B/B	Black	Gray	Gray

We see that at the 1 locus, the B allele is dominant, i.e., any mouse that has at least one B allele has black hair, but that effect is overridden by the 2 locus. Every mouse having the G allele has gray hair, regardless of which B allele they have. The effect of locus 1 is not observable unless the genotype at locus 2 is g/g. This is an example of epistasis where the effect of one locus is modified by the presence or absence of an allele at another locus. One can easily see how with phenotypes expected to be affected by several loci, epistasis can occur.

There are several methods to deal with testing for epistasis. The most common way uses standard methods to incorporate several SNPs into the same statistical model, often with terms defining the interactions between loci. The success of this method depends heavily on the number of loci being tested and the size of the sample. Putting too many terms into a single model runs the risk of overfitting the model, where the model describes random error instead of the underlying relationship between variables and usually arises when the model is excessively complex in relation to the amount of data available. The larger the sample size, the more complex a model it can support. It is easy to imagine the problem where you have a continuous phenotype and only two SNPs. When investigating a single SNP, the phenotype mean for individuals of each genotype is calculated and the three groups are compared. With two SNPs, if the means for each combination of genotypes are compared, there are now nine groups to compare. The sample size needed to adequately compare nine means is much larger than that needed to compare three means. Additionally, if the minor allele frequencies are low, a much larger sample size is needed to see any subject's homozygous mutant for both SNPs.

A more recent method of combining SNPs into a single model is called the genetic risk score [16], which combines data from several SNPs into a composite value.

There are a few ways to derive this genetic risk score. One can test each phenotype/single SNP association and use the regression coefficients from those models to weight a final derived composite score for each individual. This score is then used as the independent variable in a model with the phenotype of interest. Horne et al. [16] describe this method in detail. Another method simply looks at each phenotype/SNP association and determines which allele is the "risk" allele. The number of risk alleles for all SNPs is then counted for each individual and phenotype of interest is tested against this genotype score [16, 17]. For example, when testing four SNPs against body mass index (BMI), each SNP has a risk allele; therefore, each individual can have 0, 1, 2, 3, or 4 total risk alleles. Mean BMI is compared between these five groups to determine if there is a trend with BMI increasing and an increasing number of risk alleles.

Sample Selection Considerations

Any sample population can be affected by confounding measured or unmeasured factors. When testing hypotheses using genetic data, another form of confounding by ethnicity or population stratification can occur [18]. Population stratification arises when there is a systematic difference in allele frequencies between subpopulations within the large sample due to differing ancestry [19, 20]. Allele frequencies can vary widely between ethnic groups that can be problematic in association studies. If an allele is more frequent in one ethnic group and the sample has many individuals of that group, the strength of the association could be positively biased. Alternatively, if a true relationship between an allele and an outcome exists only in one ethnic group, but the sample has few individuals of that group, the association, if present, may not be detected.

There are several methods to deal with population stratification in a sample population of unrelated individuals; the two most common being genomic control and structured association. Genomic control uses a nonparametric method to control for any influence of shared genetic background by adjusting the underlying distribution used for determining statistical significance and is described in detail in Devlin and Roeder [21]. Structured association uses unlinked genetic markers to infer details of the population structure and to estimate the ancestry of individuals in the sample [15]. This information, described in Pritchard and Rosenberg [20], is then controlled for in the analysis to remove any possible effects of population stratification.

Candidate SNP Association Studies

One of the most important steps when analyzing SNP data is adequately defining the phenotype or outcome of interest whose form determines the statistical methods that will be used (Table 2.1). The website titled a New View of Statistics (http://www.sportsci.org/resource/stats/) [22] developed by Hopkins details all of the

Table 2.1 Typical statistical methods for the analysis of SNP data

Dependent variable (outcome/phenotype) data type	Independent variable (SNP) data type	Statistical test/model	Information gained
Categorical	Categorical	χ (chi) square test	Detects significant association between outcome and SNP
Categorical with two levels (dichotomous)	Categorical	Logistic regression	Detects significant association between outcome and each level of SNP; provides OR and 95% CI to assess the strength and direction for association
Categorical with three or more ordered levels	Categorical	Ordinal regression	Same as logistic regression
Categorical with three or more unordered levels	Categorical	Nominal regression	Same as logistic regression
Continuous	Categorical with two levels	Independent *t*-test	Detects a significant difference in means between levels of the categorical variable
Continuous	Categorical with three or more levels	ANOVA	Detects a significant difference in means between levels of the categorical variable
Continuous	Categorical with three or more levels	ANCOVA	Detects a significant difference in means between levels of the categorical variable; allows the addition of covariates
Continuous with repeated measures	Categorical with two levels	Paired *t*-test	Detects a significant difference in means between levels of the categorical variable; allows the use of phenotypes measured at multiple times
Continuous with repeated measures	Categorical with three or more levels	Repeated measures ANOVA/ANCOVA	Detects a significant difference in means between levels of the categorical variable; allows the use of phenotypes measured at multiple times
Continuous	Categorical	MANOVA	Detects significant differences in means between levels of the categorical variable; additionally allows testing of interaction between multiple dependent variables
Continuous but nonnormally distributed	Categorical with two levels	Wilcoxon sign rank (Mann–Whitney U test)	Detects a significant difference in medians between levels of the categorical variable
Continuous but nonnormally distributed	Categorical with three or more levels	Kruskal–Wallis test	Detects a significant difference in medians between levels of the categorical variable

methods discussed here with examples taken from exercise and sports science and is a valuable resource for analysis (see Appendix).

The two major types of phenotypes are those which are categorical and those which are continuous. Further, continuous data can be categorized into groups to be analyzed as categorical data, but this practice should be discouraged. Categorizing continuous data can cause problems beyond the obvious loss of information, such as a decrease in statistical power, differing conclusions depending on the categorization scheme, and the introduction of an erroneous association [23]. Therefore, it is best to treat continuous data as a continuous trait rather than categorizing it even though categorized data are sometimes more easily interpreted and better understood clinically.

The method commonly used to analyze a categorical outcome is the Chi square (χ^2) test [22]. This test indicates if there is a statistically significant association between the categorical outcome and the SNP of interest by comparing the distribution of genotype frequencies between those in different groups of the category. While this test is simple to perform, it gives only a minimum amount of information, namely whether an association exists or not. It does not give detailed information on the strength of the association, or if one particular genotype shows a stronger association with the outcome than the others. In order to determine that information, the analysis must move to a regression model; logistic regression if the categorical outcome is dichotomous or takes on only two possible values, ordinal logistic regression if the outcome takes on three or more ordered levels, and nominal logistic regression if the outcome takes on three or more unordered levels. These regression models allow measurement not only of the strength of the association, but also to specifically test each of the genotypes or alleles against one another. Each logistic regression model reports an odds ratio (OR) for each level of the factor (normally a genotype or an allele) in reference to the baseline level (usually the homozygous wild-type subjects or those with the most common genotype), which quantifies the risk (and the direction) associated with the outcome given each level of the factor. Each OR has a p value and a 95% confidence interval, allowing one to determine the strength of the association. An OR of 1.00 denotes no excess risk associated with that factor, an OR > 1.00 an excess risk, and an OR < 1.00 a decreased risk. If comparing genotypes, each can be tested against the other to determine if an excess risk is associated with any genotype. For example, a logistic regression model comparing a SNP in the *FTO* gene (rs3751812) with type 2 diabetes mellitus (T2D) showed an OR = 1.16 with a $p = 0.008$ and a 95% confidence interval of 1.04–1.30 [24]. In this instance, the number of individuals with and without the mutant allele was compared to the number of individuals with and without T2D. The OR = 1.16 showed the odds of having T2D with the mutant allele was 1.16 times greater than the odds of having T2D with the common allele. The 95% confidence interval shows the amount of uncertainty in the calculation and that the true OR is between 1.04 and 1.30.

The analysis of continuous outcomes must first start with a determination of normality. The normality of each outcome, or quantitative trait as continuous outcomes are often referred to, must be tested to determine whether parametric or

nonparametric tests are appropriate. Although parametric statistical tests (e.g., *t*-tests and analysis of variance [ANOVA]) are better known than their nonparametric counterpart tests (e.g., Wilcoxon and Kruskal–Wallis), using a parametric test on data that violates the assumptions of normality (and variance homogeneity) can lead to biased error rates severe enough to incorrectly reject the null hypothesis [25, 26]. Therefore, assessing normality, graphically using a histogram, and/or using a normality test such as the Anderson-Darling or the Shapiro–Wilk test is one of the most important initial steps in the analysis of data. If the phenotype is not normally distributed, a variety of data transformations such as log, square root, and inverse are available, but the normality of the transformed data should be assessed to determine if the transformation provided a normally distributed phenotype.

The simplest analysis of normally distributed continuous data is done using a *t*-test for comparing a phenotype between two genotypes and ANOVA for comparing three genotypes. These statistical tests will determine if there is a statistically significant difference in means between genotypes. If a significant difference is found with an ANOVA, an additional step of testing multiple comparisons is necessary. ANOVA models yield an F statistic denoting at least two of the means being compared are significantly different, but it does not define which of the means are different. To determine which means are different, post hoc pair-wise statistical tests are performed between each genotype, and the resulting *p* values are adjusted for multiple comparisons using one of the several methods available (i.e., Bonferroni, Tukey, LSD, or Sidak).

If there are other measurements, typically demographic characteristics such as age and gender, that can affect the phenotype or the relationship between the phenotype and the SNP, they can be incorporated into the analysis as covariates. Analysis of covariance (ANCOVA) is used to test differences between means, while simultaneously adjusting for or removing the variance attributable to the covariates.

If the outcome or phenotype includes repeated measurements made over time, a paired *t*-test (if comparing only two groups) or a repeated measures ANOVA or ANCOVA (if comparing three or more groups) is used to test for differences between groups. These methods allow partitioning and accounting for between-subjects and within-subjects variance and to test for significance of those effects. They also allow a test for interactions between groups (usually genotype) and the repeated measurement to see if the change in measurement is different between groups.

Lastly, nonnormally distributed continuous data must be analyzed using nonparametric tests. There are several nonparametric tests available including the Wilcoxon rank-sum (also known as the Mann–Whitney U test), which is analogous to an independent *t*-test used to compare a continuous outcome between two independent groups. The Wilcoxon sign rank test is used in place of a paired *t*-test for comparing a continuous outcome between two paired groups. When comparing three independent groups, the Kruskal–Wallis test is used in place of ANOVA. More complicated analyses with repeated measures data in more than three groups or with the inclusion of covariates can be done by taking the continuous data, ranking the values, and using parametric tests on the ranks rather than the raw data values.

Genome-Wide Association Studies (GWAS)

GWAS are defined by the National Institutes of Health as a study of common genetic variation across the entire genome designed to identify genetic associations with observable traits [5] and have become more and more common as the technology to perform them has evolved and improved. Over 100 loci for more than 40 common diseases have been identified though GWAS in just the last few years and that number is growing [5]. The National Human Genome Research Institute (NHGRI) produces a searchable catalog of all published GWAS freely available for download from http://www.genome.gov/gwastudies/ (see Appendix).

GWAS can be described as unbiased but comprehensive association studies investigating most of the genome for causal variants [19]. An outcome, either categorical or continuous, is tested for association with an enormous number (currently one million) of SNPs located throughout the genome. GWAS is an especially important development in the study of complex diseases, which most likely involve many loci, each of which contributes minimally to the effect size [27]. Traditional family-based linkage studies have been successful in identifying genes with large effect sizes, but not multiple genes with small effects. GWAS offer an unprecedented opportunity to study the genetic unpinning of common and complex diseases by interrogating the entire genome in unrelated individuals without the need for prior hypotheses, but this revolutionary technology has some limitations that must be understood and addressed.

The greatest benefit of a GWAS is that no prior assumption of the location of a causal variant is needed. Therefore, GWAS can be performed without having any idea of where a causal variant may be located or even what the function of the causal variant may be [19, 27]. This approach differs greatly from candidate gene association studies where the loci investigated are chosen based on prior knowledge of biological meaning and/or location. When the biological underpinnings of a disease are completely unknown or when the location of the possible gene involved is unknown, candidate gene association studies are not very efficient [27] because a prior knowledge of the loci being studied is needed. GWAS, on the other hand, can scan the entire genome to find loci that are possibly associated with the outcome. There is no prior guessing as to what genes are involved in the outcome or where they are located in GWAS.

There are several things to be considered when performing a GWAS and when interpreting the results of a published GWAS report. These considerations include, among other things, the study design, selection of study subjects, sample size, quality control, analysis of the data, and replication. An excellent review published by the NCI-NHGRI Working Group on Replication in Association Studies in 2007 describes in detail those factors to consider when reading or performing a GWAS [28].

One of the first considerations in a GWAS is the study design and what type of subjects will be used. There are a few study designs that can be used for GWAS, the most common of which is the case/control design where the outcome is a dichotomous yes or no answer, but cohort designs where the outcome is a quantitative trait

are beginning to be published. Both designs have advantages and disadvantages. Case control samples are usually drawn from the same population, decreasing the chance of bias or differences between the groups unrelated to the factor(s) defining cases, but the classification of cases must be very rigorous. Misclassification of cases, and less importantly controls, can markedly reduce the power to detect associations [5]. Appropriate control of possible confounders must be done to avoid the risk of identifying associations not with the outcome of interest, but with the confounding factor instead [5]. Cohort samples have a better chance of homogeneity of participants, decreasing the chance of bias, but usually require a larger sample size, especially if the expected effect size is small [5]. Both types of samples are vulnerable to population stratification, and it is recommended GWAS assess it. A more recent development is the multistage design where a smaller number of individuals are subjected to genotyping the entire genome; only those SNPs found to be significant or promising are then genotyped in successively larger samples.

The NCI-NHGRI Working Group on Replication in Association Studies recommends the sample size for a GWAS should be sufficient to detect a modest effect size and be "suitably large," but does not specifically define the term [28]. Simulation studies [29] have shown the power to detect a moderate effect size of 1.3 is nearly zero with a sample size of 1,000 cases and 1,000 controls and only increases to ~50% with a sample three times larger. A list of GWAS (http://www.genome.gov/gwastudies/) (see Appendix) published from November 2008 to November 2009 shows sample sizes ranging from the smallest at 60 to the largest at 81,000 with a wide range of effect sizes. To help with designing a GWAS and estimate the power of a proposed study given a particular sample size and budget, a module for the R package (a free software environment for statistical computing and graphics – http://www.r-project.org/) is available from http://www.stats.ox.ac.uk/~marchini/#software (see Appendix).

Quality control of genotyping is essential to any association study, including GWAS. Genotyping errors can cause spurious results, especially if the errors occur differentially between the cases and controls in a case control study [5]. Therefore, several quality control features should be present including the SNP genotyping call rate, the SNP minor allele frequencies, the results of HWE, and concordance rates in duplicate samples. NCI-NHGRI also recommends those SNPs shown to have significant associations be genotyped on a different platform to confirm results and known associations be verified in the sample under study [28].

Many of the issues arising from GWAS stem from the data analysis. The problem of multiple testing described in an earlier section becomes a critical issue when testing upwards of one million SNPs in a single experiment [30]. There are methods to adjust the resulting p values to deal with the issue of multiple testing and reduce the false positive rate. The most commonly applied is the Bonferroni approach where the critical p value to determine significance is divided by the number of tests done (i.e., if the nominal critical p value is 0.05, the new critical value for significance with 1,000,000 SNPs is 5×10^{-8}). This adjustment has been criticized for being overly conservative because it assumes each statistical test on each SNP is independent which is not true due to extensive LD between SNPs within the genome [5].

Other methods are available, but are not as widely used and require more computation beyond the amount necessary to calculate the *p* values themselves.

The last consideration for GWAS is replication. Any SNP association initially found to be significant should ideally be replicated in another independent sample [20]. This replication can be built in to the initial study through a multistage design or can be made in an entirely different sample with similar characteristics. However, if the finding from the initial GWAS is a false positive, it will not be detected in a follow-up replication study.

The analysis of data from a GWAS is computationally intensive and requires a computer program designed specifically for its analysis. There are several commercially available programs including Partek Genomic Suite (http://www.partek.com/) and Golden Helix (http://www.goldenhelix.com/). There are also several freely downloadable programs available. The most commonly used is PLINK (available at http://pngu.mgh.harvard.edu/~purcell/plink/) and a list of other programs is given at http://www.stats.ox.ac.uk/~marchini/software/gwas/gwas.html (see Appendix).

Using GWAS as a Tool for Exercise Science

GWAS are valuable tools for the exercise scientist and clinician. A GWAS is an unbiased method to ascertain genetic variation (in the form of SNPs) in the human genome, and the tools to analyze these scans are available (discussed previously). Never before has one been able to scan the genomic landscape with the precision of a GWAS. This development has resulted in the identification of new genes or regions of the genome that may have some biological significance for a particular phenotype such as skeletal muscle mass.

The statistical challenges of a GWAS are many, but it will be an important method that will be used for the foreseeable future in the study of exercise genomics. For example, many scientists feel GWAS results are genetic variants with weak effect sizes and a small fraction of heritability is explained by these variants. Heritability is the amount of phenotypic variation in a population that is explained by genetic variation among individuals. However, with the identification of new regions of the genome obtained by GWAS, new biological insights are obtained that will lead to clinical advances. This will lead to the development of new biomarkers of exercise performance and health fitness-related phenotypes which can be used to track the positive effects of exercise. For example, there are increases in the phosphorylation of p70-S6 kinase (S6K1) after resistance exercise [31]. GWAS might identify new biomarkers that have pathways that have never been implicated using in vivo models of resistance training. Understanding how these technologies will be used is a question that remains unanswered and will require thought and deliberation from exercise scientists choosing to undertake this work. This next section of the chapter will focus on the use of GWAS as a tool in exercise science including what has been done in the field thus far. Finally, the next steps for GWAS will be discussed.

With the publication of the draft sequences of the human genome project in 2001 [32, 33], the field of genomics exploded. Now, almost the entire sequence of the human genome is completed [34] and is in searchable online databases (http://www.ncbi.nlm.nih.gov/ or http://genome.ucsc.edu/) (see Appendix). Even with the completed human genome sequence, the number of genes remains a mystery. An analysis by scientists at Ohio State University suggested the number of genes was between 65,000 and 75,000 [35], while another study published in *Cell* in 2001 predicted a total of 42,000 genes [36]. However, a recent paper found there to be only 20,500 genes [37]. The number of genes was lowered based on a rejection of open reading frames that are not conserved between species. The final number of human genes will not be known for many years and will differ based on different definitions of a gene.

Why is the number of genes important for GWAS? GWAS is designed as a method to survey the entire genome using SNPs distributed at certain intervals of base pair (bp) length in an attempt to capture the genetic diversity of the individual [19]. Therefore, understanding the coding and noncoding regions of the genome is important. The goal of the Human Genome Project was to provide scientists with powerful new tools to help them clear the research hurdles that now keep them from understanding the molecular essence of other tragic and devastating illnesses, such as schizophrenia, alcoholism, Alzheimer's disease, and manic depression.

There is predicted to be ten million sites (SNPs) that are variable (at least 1%) in human populations (about 1 SNP per 300 bp) [38]. With this information, identifying SNPs associated with a phenotype such as skeletal muscle strength should be straightforward and involve genotyping the ten million sites to discover frequency differences for the SNPs between the individuals who are stronger vs. those who are not as strong. However, there are cost issues and structural polymorphism variants that would not be identified (i.e., insertions, deletions, and inversions) using GWAS which only examines SNPs. Therefore, methods to complete a GWAS remain to be developed.

There have been numerous studies showing that individuals who carry a particular SNP allele at one site often predictably carry specific alleles at other nearby variant sites. This correlation is known as LD; a particular combination of alleles along a chromosome is termed a haplotype. LD exists because of the shared ancestry of contemporary chromosomes [38]. The International HapMap project (http://hapmap.ncbi.nlm.nih.gov) (see Appendix) was developed to find these genome-wide database of patterns of human genetic variation [38]. The HapMap project was a logical next step in defining the human genome and was started in 2003 [38].

Four different populations with African, Asian, and European ancestry were used in the completion of the HapMap (http://hapmap.ncbi.nlm.nih.gov) project that sought to determine the frequencies and patterns of association among roughly three million common SNPs for use in GWAS. The website can be used with genomic positions or gene names and will show the SNPs that are in LD (or linked). This website can be used to reduce the number of SNPs used for a GWAS study because redundancy is eliminated.

The HapMap project ushered in new technological advances that made GWAS a possibility with cost-efficient methods to genotype up to one million SNPs.

In addition, new insights into the evolutionary pressures on the human genome began to be understood. For example, functional portions of the genome (those that code for proteins) are under selective pressure to contain less polymorphisms [39]. Changes in the amino acid sequence of a protein by a SNP may render the protein useless, and thus there would be selective pressure by evolution to remove that SNP from the gene pool. Finally, understanding the function of genetic variation (which will be discussed later in this chapter) has become a priority.

GWAS is a fantastic tool that has been used to find strong associations between a specific chromosomal locus and complex human diseases such as T2D (please see Chap. 4 for additional information on T2D). The single largest effect size for T2D identified to date resides within the transcription factor 7-like 2 (*TCF7L2*) gene [40]. This gene has been replicated in numerous GWAS [41–43]. In addition, there have been numerous loci identified that are associated with lipids such as variants within or in close proximity to the Sortilin 1 (*SORT1*) and Cadherin, EGF LAG seven-pass G-type receptor 2 (*CELSR2*) genes [17, 44, 45].

GWAS is in its infancy for understanding exercise genomics. Only one such study has been published on a skeletal muscle phenotype, lean body mass (LBM). Liu et al. [46] found two SNPs, rs16892496 ($p = 7.55 \times 10^{-8}$) and rs7832552 ($p = 7.58 \times 10^{-8}$), within the thyrotropin-releasing hormone receptor (*TRHR*) gene were significantly associated with LBM after testing 379,319 eligible SNPs in 1,000 unrelated US Whites. Subjects carrying unfavorable genotypes at rs16892496 and rs7832552 had, on average, 2.70 and 2.55 kg lower LBM, respectively, compared to those with alternative genotypes. As is important for a GWAS, the results were replicated in three independent samples: (1) 1,488 unrelated US Whites, (2) 2,955 Chinese unrelated subjects, and (3) 593 nuclear families comprising 1,972 US Whites [46]. The meta-analyses of the GWA scan and the replication studies yielded p values $= 5.53 \times 10^{-9}$ for rs16892496 and 3.88×10^{-10} for rs7832552. This functional data point to the *TRHR* gene as a component of the mechanism for LBM.

The *TRHR* gene encodes the TRHR, which belongs to the G protein-coupled receptor 1 family. Thyrotropic-releasing hormone (TRH) is a tripeptide (Glu-His-Pro) hormone secreted by the hypothalamus. TRH binds to TRHR on the surface of pituitary thyrotrophs. The primary consequence of the TRH:TRHR binding is activation of the inositol phospholipid-calcium-protein kinase C transduction pathway that stimulates secretion of thyroid-stimulating hormone (TSH) and prolactin. The TSH response to TRHR is the first step in the hormonal cascade of hypothalamic-pituitary-thyroid axis (HPTA) that eventually leads to the release of thyroxin, which is important in the development of vertebrate skeletal muscle [47]. Additionally, thyroxin is required for the full anabolic action of the growth hormone-insulin-like growth factor-I (GH-IGF1) axis that is central in its role in muscle protein balance and adaptive changes to resistance training [46]. Results of this study, together with the functional relevance of TRHR in muscle metabolism, support the *TRHR* gene as an important gene for LBM variation. This type of work is an excellent model for the type of work that needs to be performed to examine the influence of genetic variation on health fitness and exercise performance phenotypes in exercise science.

GWAS as a Tool to Uncover Functional Gene Variants

One of the most exciting areas of study within the realm of GWAS is the uncovering of functional genetic variants that may control gene expression. Gene expression or a change in the level of mRNA is the underlying force for cellular phenotypes including muscle growth [48, 49]. For example, when a skeletal muscle undergoes contraction, IGF1 is secreted [49]. Gene expression has been used to understand the pathways that underlie the response of skeletal muscle to resistance training [50]. The understanding of how gene expression regulates key biological functions is not well understood and is just beginning to be explored. The technical development of microarrays as a biological tool has allowed the measure of gene expression on a global scale, which has changed the face of molecular biology [51]. Gene expression is used to discover pathways that drive diseases such as muscular dystrophy [52] and/or the response of skeletal muscle to resistance training [50]. This development is now coupled to GWAS to identify individual differences in the quantitative levels of gene expression (i.e., expression quantitative loci [eQTLs]). In exercise science, this will be a key to the development of new pathways that affect skeletal muscle response or aerobic response to training.

Gene changes that are identified using gene expression profiling can be used in conjunction with GWAS data to map genetic variants that have an effect on gene expression. This will be valuable for the discovery of new genes and the genetic variants that influence gene expression involved in the response of the skeletal muscle to resistance training. But this can be used with any type of exercise in any type of tissue. For example, the brain of a mouse undergoing aerobic training could be expression profiled and these data analyzed with genetic variants to find SNPs that control brain response to aerobic training. In addition, the discovery of new genetic variants and the genes that they exert control over may solidify the current pathways present in the literature. If a genetic variant that elevates the expression of *IGF1* is identified using GWAS and whole genome expression profiling, it is possible that individuals with that variant may have an improved response to some types of resistance training. These new tools will lead to the development of an arsenal of new genes for further study by matching gene expression data with GWAS data. In addition, these new pathways maybe specific to skeletal muscle and will need to be explored in the other tissues in the human body.

The Future?

Throughout this chapter, we have included comment on our vision for the future regarding statistical and methodological considerations. In summary, as technology advances, the data management and statistical tools to effectively analyze the resulting data will evolve. Platforms to genotype larger sets of SNPs are being developed, and the methods to deal massive amounts of data will be needed. Along with SNP data, data from other sources such as gene expression and proteomics will have

to be integrated in order to get a more complete picture of biological processes. The development of next-generation sequencing will require statistical methods to deal with the data it produces. All of these are areas of active research that are driven by the increasing technical ability to produce more detailed biological information. The exercise scientist doing genomics research will need to become an expert in understanding complicated genetics and the statistical methods to process the data.

Practical Applications

We have commented on practical applications throughout this chapter. Briefly, all of the statistic methods discussed in this chapter can and are being used by exercise scientists. The advent of new genotyping methods involving larger data sets and methods to interrogate more genetic variation will require new methods of data handling. In addition, the methods in this chapter will be used to examine minor statistical effects from genetic variation.

References

1. Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, et al. Detection of genotyping errors by Hardy–Weinberg equilibrium testing. Eur J Hum Genet. 2004;12(5):395–9.
2. Trikalinos TA, Salanti G, Khoury MJ, Ioannidis JP. Impact of violations and deviations in Hardy–Weinberg equilibrium on postulated gene-disease associations. Am J Epidemiol. 2006;163(4):300–9.
3. Stern C. The Hardy–Weinberg law. Science. 1943;97(2510):137–8.
4. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science. 2002;296(5576):2225–9.
5. Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008;299(11):1335–44.
6. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. Am J Hum Genet. 2001;69(1):1–14.
7. Weiss KM, Clark AG. Linkage disequilibrium and the mapping of complex human traits. Trends Genet. 2002;18(1):19–24.
8. Wallace C, Dobson RJ, Munroe PB, Caulfield MJ. Information capture using SNPs from HapMap and whole-genome chips differs in a sample of inflammatory and cardiovascular gene-centric regions from genome-wide estimates. Genome Res. 2007;17(11):1596–602.
9. Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics. 1995;29(2):311–22.
10. Qin S. Linkage disequilibrium: Biostat830: advanced topics in biostatistics lecture notes. http://www.sph.umich.edu/~qin/biostat830/LDnote.doc. Accessed 9 Nov 2009.
11. Wall JD, Pritchard JK. Haplotype blocks and linkage disequilibrium in the human genome. Nat Rev Genet. 2003;4(8):587–97.
12. Pistilli EE, Gordish-Dressman H, Seip RL, Devaney JM, Thompson PD, Price TB, et al. Resistin polymorphisms are associated with muscle, bone, and fat phenotypes in white men and women. Obesity (Silver Spring). 2007;15(2):392–402.
13. Orkunoglu-Suer FE, Gordish-Dressman H, Clarkson PM, Thompson PD, Angelopoulos TJ, Gordon PM, et al. INSIG2 gene polymorphism is associated with increased subcutaneous fat in women and poor response to resistance training in men. BMC Med Genet. 2008;9:117.

14. Zheng G, Joo J, Zhang C, Geller NL. Testing association for markers on the X chromosome. Genet Epidemiol. 2007;31(8):834–43.
15. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet. 2002;11(20):2463–8.
16. Horne BD, Anderson JL, Carlquist JF, Muhlestein JB, Renlund DG, Bair TL, et al. Generating genetic risk scores from intermediate phenotypes for use in association studies of clinically significant endpoints. Ann Hum Genet. 2005;69(Pt 2):176–86.
17. Kathiresan S, Melander O, Anevski D, Guiducci C, Burtt NP, Roos C, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med. 2008; 358(12):1240–9.
18. Thomas DC. Statistical methods in genetic epidemiology. Oxford: Oxford University Press; 2004.
19. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet. 2005;6(2):109–18.
20. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999;65(1):220–8.
21. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55(4): 997–1004.
22. Hopkins WG. A new view of statistics. http://www.sportsci.org/resource/stats/index.html2010. Accessed 4 November 2009.
23. van Walraven C, Hart RG. Leave 'em alone – why continuous variables should be analyzed as such. Neuroepidemiology. 2008;30(3):138–9.
24. Ng MC, Tam CH, So WY, Ho JSK, Chan AW, Lee HM, et al. Implication of genetic variants near NEGR1, SEC16B, TMEM18, ETV5/DGKG, GNPDA2, LIN7C/BDNF, MTCH2, BCDIN3D/FAIM2, SH2B1, FTO, MC4R, and KCTD15 with obesity and type 2 diabetes in 7705 Chinese. J Clin Endocrinol Metab. 2010;95(5):2418–25.
25. Cribbie RA, Keselman HJ. The effects of nonnormality on parametric, nonparametric, and model comparison approaches to pairwise comparisons. Educ Psychol Meas. 2003;63(4):615–35.
26. Pappas PA, DePuy V. An overview of non-parametric tests in SAS®: when, why, and how. http://analytics.ncsu.edu/sesug/2004/TU04-Pappas.pdf2010. Accessed 6 November 2009.
27. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005;6(2):95–108.
28. NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, et al. Replicating genotype-phenotype associations. Nature. 2007;447(7145):655–60.
29. Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet. 2009;5(5):e1000477.
30. Hunter DJ, Kraft P. Drinking from the fire hose – statistical issues in genomewide association studies. N Engl J Med. 2007;357(5):436–9.
31. Nader GA, Esser KA. Intracellular signaling specificity in skeletal muscle in response to different modes of exercise. J Appl Physiol. 2001;90(5):1936–42.
32. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860–921.
33. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001;291(5507):1304–51.
34. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004;431(7011):931–45.
35. Wright FA, Lemon WJ, Zhao WD, Sears R, Zhuo D, Wang J, et al. A draft annotation and overview of the human genome. Genome Biol. 2001;2(7):1–18.
36. Hogenesch JB, Ching KA, Batalov S, Su AI, Walker JR, Zhou Y, et al. A comparison of the celera and ensembl predicted gene sets reveals little overlap in novel genes. Cell. 2001;106(4):413–5.

37. Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, et al. Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci U S A. 2007; 104(49):19428–33.
38. International HapMap Consortium. The international HapMap project. Nature. 2003; 426(6968):789–96.
39. Davidson S, Starkey A, MacKenzie A. Evidence of uneven selective pressure on different subsets of the conserved human genome; implications for the significance of intronic and intergenic DNA. BMC Genomics. 2009;10:614.
40. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881–5.
41. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PW, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007;316(5829):1331–6.
42. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet. 2008;40(5):638–45.
43. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007;316(5829):1336–41.
44. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet. 2008;40(2):189–97.
45. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet. 2008;40(2):161–9.
46. Liu XG, Tan LJ, Lei SF, Liu YJ, Shen H, Wang L, et al. Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. Am J Hum Genet. 2009;84(3):418–23.
47. Larsson L, Li X, Teresi A, Salviati G. Effects of thyroid hormone on fast- and slow-twitch skeletal muscles in young and old rats. J Physiol. 1994;481(Pt 1):149–61.
48. Coffey VG, Hawley JA. The molecular bases of training adaptation. Sports Med. 2007;37(9):737–63.
49. Glass DJ. Molecular mechanisms modulating muscle mass. Trends Mol Med. 2003;9(8): 344–50.
50. Chen YW, Nader GA, Baar KR, Fedele MJ, Hoffman EP, Esser KA. Response of rat muscle to acute resistance exercise defined by transcriptional and translational profiling. J Physiol. 2002;545(Pt 1):27–41.
51. Mahoney DJ, Tarnopolsky MA. Understanding skeletal muscle adaptation to exercise training in humans: contributions from microarray studies. Phys Med Rehabil Clin N Am. 2005;16(4):859–73, vii.
52. Hoffman EP, DuBois DC, Hoffman RI, Almon RR. Expression profiling and pharmacogenomics of muscle and muscle disease. Curr Opin Pharmacol. 2003;3(3):309–16.

http://www.springer.com/978-1-60761-354-1

Exercise Genomics
Pescatello, L.S.; Roth, S.M. (Eds.)
2011, XX, 268 p., Hardcover
ISBN: 978-1-60761-354-1
A product of Humana Press