Austin Biometrics and Biostatistics

Open Access

Austin Publishing Group

Review Article

Region-Based Tests for Association Analysis of Rare Variants

Badri Padhukasahasram*
Center for Health Policy and Health Services Research, Henry Ford Health System, USA

***Corresponding author:** Badri Padhukasahasram, Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, Michigan, 1 Ford Place 3A, Henry Ford Health System, Detroit, Michigan, USA

Received: January 05, 2015; **Accepted:** February 18, 2015; **Published:** March 02, 2015

Abstract

Despite numerous discoveries based on genome wide association studies of common variants, the heritability of most complex traits remains largely unexplained. Rare variants may play a significant role in disease risk and phenotypic variation. Such variants are known to be associated with mendelian disorders and rare forms of common diseases. They are also known to be associated with complex diseases. Dramatic advances in DNA sequencing technologies have enabled a more comprehensive evaluation of the full spectrum of genetic variation and now enable us to evaluate the role of low frequency and rare variation in complex traits. In this review, I provide an overview of the various methods that are available for testing simultaneous association of multiple rare variants with disease or any other phenotypes in the context of sequencing based association studies. The tests focus on rare variation from a particular genomic region such as a gene and its surrounding regions. I discuss the basic underlying ideas behind many currently available approaches for region-based association testing of rare variants as well as their advantages and limitations.

Keywords: Rare variants; Burden tests; Variance-component tests; Ominbus tests; Exponential combination tests; Power; Regression; Gene-based association; Region-based association

Introduction

Common Disease Common Variant hypothesis (CDCV) has been a main driver of numerous Genome Wide Association Studies (GWASs) in the last decade [1]. The CDCV hypothesis asserts that common diseases are caused by common variants (frequency > 5%) with low to modest effects [2-5]. The studies of such variants have led to a large number of discoveries [6] and have yielded valuable insights into the genetic basis of complex phenotypes [7-12]. Despite these discoveries, for most complex traits, a large fraction of the genetic contribution as would be expected from heritability estimates (e.g. from twin studies) remains unexplained. For example, for Type 2 Diabetes and Crohn's disease even with sample sizes of association studies reaching a range of > 100,000, all the current discoveries taken together can only explain ~11% and ~23% of the respectively of the heritability. This so called "missing heritability" problem has received a great deal of attention in the recent times and several explanations [13,14] have been formulated to account for the rest of the genetic contribution to disease and complex traits. If heritability estimates available are accurate, then the missing genetic contribution could be in the form of variation that has not been as extensively investigated as common variation. Because of the CDCV hypothesis, GWASs have focused on the identification of common variants with Minor Allele Frequency (MAF) larger than 5%; however, the rest of the frequency spectrum may contain additional trait-associated variation (e.g. low frequency variants MAF in range 1-5% and rare variation MAF < 1%).

In particular, rare variants can play a significant role in disease risk and phenotypic variation. Such variants are known to be associated with mendelian disorders and rare forms of common diseases [15]. There is also a growing body of evidence that rare variants are associated with complex phenotypes [16-22]. Dramatic advances in DNA sequencing technologies now enabled us to evaluate the role of low frequency and rare variation in complex traits [23-25] High-throughput sequencing technologies can generate billions of short reads across the genome at a reasonable cost and have made whole-exome and whole-genome sequencing studies feasible. Improved sequencing technologies as well as rare-variant genotyping chips [26] have led to genome wide scans for detecting rare variant associations. These are also referred to as Rare Variants Association Studies (RVAS). In [27], sequencing of whole exomes was carried out in 3,734 individuals to test for associations with plasma triglyceride levels. Carriers of rare loss-of-function mutations in the APOC3 gene were found to have 39 percent lower triglyceride levels than non-carriers, as well as better cholesterol levels. In [28], analysis of rare coding variation in 3,871 autism cases and 9,937 ancestry-matched or parental controls revealed 22 autosomal genes. In [29], researchers sequenced the exomes of 2,536 cases with schizophrenia and 2,543 unrelated controls. Schizophrenia cases had a significantly higher rate of rare disruptive mutations in protein-coding schizophrenia candidate genes.

In contrast to common variants, the detection and subsequent association testing with rare variants presents many challenges. Firstly, large sample sizes are needed simply to observe a rare variant in the sample. Secondly, the standard single-variant association tests designed for common variants are underpowered when used for finding rare variant associations. Because deep whole genome sequencing of large sample sizes is currently cost prohibitive, the first issue can be solved by alternate strategies such as targeted sequencing

Citation: Padhukasahasram B. Region-Based Tests for Association Analysis of Rare Variants. Austin Biom and Biostat. 2015;2(1): 1012.

[30], exome sequencing [31], extreme-phenotype sampling [32-35] and low-coverage sequencing [36,37]. To address the power issue, numerous region-based multi-marker tests have been proposed in the last several years [38,39]. In this review, I provide an overview of the various methods that are available for testing simultaneous association of multiple rare variants with disease or any other phenotypes in the context of sequencing based association studies. The tests focus on rare variation from a particular genomic region such as a gene and its surrounding regions. I discuss the basic underlying ideas behind many currently available approaches for region-based association testing of rare variants as well as their advantages and limitations.

Methods for association analysis of rare variants

In the classical single-variant association testing, linear or logistic regression is used for association testing and a genome wide p value threshold of 5×10^{-8} is used to account for multiple testing correction (1 million independent tests) [40]. Regression-based approaches allow us to easily adjust for covariates. For the same effect size, the power to detect association with a rare variant is expected to be smaller than for common variants [39]. The sample size needed to achieve over 80% power with rare variants is at least an order magnitude higher than common variants. Furthermore, because the total number of rare variants across the genome is also larger than common variants, correction for multiple testing will further reduce power in this case. Region-based tests of association seek to aggregate cumulative effects of multiple genetic variants in a gene or region instead of testing each variant individually. When many variants from a relevant gene or genomic region are associated with a complex trait, they may increase the power to detect such associations. Instead of testing millions of rare variants, we can test ~20,000 or so gene regions and this can help reduce the multiple testing burdens. Methods for rare variant association analysis can be classified into 4 major categories: burden tests, variance component tests, combined burden and variance-component tests and the exponential-combination test.

Burden tests

The main idea behind burden tests is to collapse information for multiple genetic variants into a single variable and test for associations between this variable and disease status [41-46]. There are many ways to combine the information from multiple genetic variants into a single score such as counting the number of minor alleles for all variants and weighting them to get a composite score. The weights can be based on minor allele frequency as well as functional information based on where a particular variant is located in the genome. These different methods are based on different assumptions about disease mechanism. In general, burden tests make strong assumptions that all the variants in a set are causal and have same direction and effect size. When a large proportion of variants are indeed causal and have same direction of effect, such tests can be powerful. Violation of these assumptions can lead to loss of power [47-49].

Adaptive burden tests [50-55] are refinements to the original burden tests idea that allow for variants to have effects in both directions. They are more robust than original burden tests because they make fewer assumptions about the underlying genetic model at each locus. At the same time, adaptive tests based on regression are often difficult and unstable for rare variants and those that make use of permutation are computationally intensive. Han et al. [50] developed a data-adaptive sum test that first estimates the direction of effect for each variant and then uses the estimated directions to conduct a burden test. The step-up test [51] refines the procedure to use a model-selection framework that assigns zero weight when a variant is unlikely to be associated.

Variance-component tests

These types of tests use a random-effects model and construct a variance-component test that evaluates the distribution of genetic effects for a set of variants. Instead of aggregating variants, these tests evaluate the distribution of the aggregated score test statistics. The Sequence Kernel Association (SKAT) [56-59], sum of squared score test [57] and the C-alpha test [58] are all based on this principle. SKAT allows for both covariate adjustment and modeling of interactions between variants. The test statistic is a weighted sum of squares of individual score statistics and asymptotically follows a mixture chi-square distribution. The p value can be computed rapidly using analytic formulas [60,61]. Variance component tests are powerful in the presence of both phenotype-increasing and phenotype-decreasing variants as well as in cases where only a small proportion of the variants are causal. However, these are less powerful than burden tests when most variants are causal and have effects are in the same direction.

Omnibus tests

Because burden and variance component tests are complementary in terms of the scenarios in which they attain high power, it is desirable to combine these two approaches. Derkach et al. [62] use Fisher's method [63] to combine the p values of these two tests and make use of permutation to evaluate the significance of the test. Another approach is to use the data to adaptively combine the SKAT and burden test statistics. Lee et al. [64] propose a linear combination of SKAT and burden test statistics. An adaptive procedure is used to find the optimal way to combine test statistics and p values are calculated through one-dimensional numerical integration. Combined tests are attractive in practice because they do not assume a particular genetic architecture and in most situations we do not have strong priors for the underlying genetic model. However, such tests can be slightly less powerful than the previous 2 categories of tests when the assumptions underlying those tests are satisfied.

Exponential combination tests

In contrast to burden and variance component tests that use linear or quadratic combination of score statistics, this test makes use of an exponential sum of the score statistics [65]. The test statistic is developed under a Bayesian framework with a sparse alternative prior with the assumption that only one variant in a genomic region is causal. The significance of the test is determined through the use of permutations. Because the exponential function increases rapidly, the exponential-combination test can have higher power when only a small proportion of the variants are causal but becomes less powerful when moderate or large proportions of variants are causal. Because the null distribution of the test statistic is unknown, permutations are used to obtain p values, making the test computationally intensive.

Relative performance: power and type 1 error rates

Although numerous rare-variant association methods have been proposed, a comprehensive comparison of their performance in

terms of power and false positive rates had been lacking until recently. Dering et al. [66] compared 15 conceptually different rare-variant association methods using simulation data for Genetic Analysis Workshop17 [67] as well as empirical data investigating methotrexate clearance in Acute Lymphoblastic Leukemia (ALL) diseased children [68]. The results of testing these 15 approaches [42-48,50,53,54,56,69-71] indicated that unexpectedly, many of proposed rare-variant association testing approaches have substantially inflated Type 1 error rates. Specifically, only methods proposed in [47,53,54,70] had valid Type 1 error rates for all the simulation scenarios considered in that study. Among all the tests with valid false positive rate, the method proposed in [47] had the largest power in the 4 scenarios that were investigated in the simulations. Findings from the empirical dataset were consistent with the comparisons from simulation study. Both simulations and analysis of real data showed that the power of collapsing based methods heavilyrelieson the proportion of causal variants in the region of interest [72]. Furthermore, not all of these approaches allow for covariate adjustment and methods assuming only a genetic effect may be at disadvantage when phenotype is influenced by covariates.

In conclusion, the study of the association of rare variants or groups of rare-variants is likely to be a major focus of future genetic association studies as we try to better understand the genetic basis of complex traits. The function of a gene can be altered by mutations in many different positions and all of these can influence the phenotype. Genes rarely work in isolation and multiple rare variants occurring in different genes that are part of a biological pathway can together affect phenotype expression. This motivates the development of valid tests that can look at the collective association of rare (and possibly common) variation in genes and biological pathways and such tests can also enhance power as compared to single variant analyses. In general, the analysis of rare variants is complicated by low power, lack of knowledge of the underlying genetic model as well as the difficulty of calling rare genotypes. Prior information about the functional importance of a variant site as derived from computational prediction tools and biological knowledge can guide the choice of regions of interest to detect truer are variant associations. Irrespective of the method used, studies with small to moderate sample sizes are likely to suffer from lack of power. Even when sufficiently large sample sizes are available, rare-variant association testing methods that rely on permutations require huge computational effort making them less appealing in practice as compared to other valid asymptotic methods.

References

1. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012; 90: 7-24.
2. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005; 6: 95-108.
3. Iyengar SK, Elston RC. The genetic basis of complex traits: rare variants or "common gene, common disease"? Methods Mol Biol. 2007; 376: 71-84.
4. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet. 2001; 17: 502-510.
5. Smith DJ, Lusis AJ. The allelic structure of common disease. Hum Mol Genet. 2002; 11: 2455-2461.
6. Hindorff LA, Junkins HA, Mehta J, Manolio T. A Catalog of Published Genome-wide Association Studies. National Human Genome Research Institute. 2010.
7. Lee JC, Parkes M. Genome-wide association studies and Crohn's disease. Brief Funct Genomics. 2011; 10: 71-76.
8. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005; 308: 385-389.
9. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Wellcome Trust Case Control Consortium; Genetic Investigation of ANthropometric Traits Consortium. 2009.
10. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009; 41: 25-34.
11. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993; 261: 921-923.
12. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 2010; 42: 565–569.
13. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010; 11: 446-450.
14. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci USA. 2012; 109: 1193-1198.
15. Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet. 2012; 13: 135-145.
16. Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, et al. National Institute of Diabetes and Digestive KidneyDiseases Inflammatory Bowel Disease Genetics Consortium(NIDDK IBDGC); United Kingdom Inflammatory BowelDisease Genetics Consortium; International Inflammatory Bowel Disease Genetics Consortium. Deep resequencingof GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet. 2011; 43: 1066–1073.
17. Gudmundsson J, Sulem P, Gudbjartsson DF, Masson G, Agnarsson BA, Benediktsdottir KR, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. Nat Genet. 2012; 44: 1326-1329.
18. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012; 488: 96–99.
19. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science. 2004; 305: 869-872.
20. Cohen JC, Pertsemlidis A, Fahmi S, Esmail S, Vega GL, Grundy SM, et al. Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. Proc Natl Acad Sci USA. 2006; 103: 1810-1815.
21. Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, et al . Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet. 2008; 40: 592-599.
22. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science. 2009; 324: 387-389.
23. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet. 2010; 11: 415-425.
24. Nelson MR, Wegmann D, Ehm MG, Kessner D, St Jean P, Verzilli C, et al. An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. Science. 2012; 337: 100-104.
25. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491: 56-65.

26. Wagner MJ . Rare-variant genome-wide association studies: a new frontier in genetic analysis of complex traits. Pharmacogenomics. 2013; 14: 413-424.

27. TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL . Loss-of-function mutations in APOC3, triglycerides, and coronary disease. N Engl J Med. 2014; 371: 22-31.

28. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014; 515: 209-215.

29. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511: 421-427.

30. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet. 2010; 42: 684-687.

31. Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stancakova A, Stringham HM, et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat Genet. 2013; 45: 197-201.

32. Guey LT, Kravic J, Melander O, Burtt NP, Laramie JM, Lyssenko V, et al. Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. Genet Epidemiol. 2011; 35: 236-246.

33. Barnett IJ, Lee S, Lin X. Detecting rare variant effects using extreme phenotype sampling in sequencing association studies. Genet Epidemiol. 2013; 37: 142-151.

34. Li D, Lewinger JP, Gauderman WJ, Murcray CE, Conti D. Using extreme phenotype sampling to identify the rare causal variants of quantitative traits in association studies. Genet Epidemiol. 2011; 35: 790-799.

35. Emond MJ, Louie T, Emerson J, Zhao W, Mathias RA, Knowles MR, et al. National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project; Lung GO. Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic Pseudomonas aeruginosa infection in cystic fibrosis. Nat Genet. 2012; 44: 886–889.

36. Li Y, Sidore C, Kang HM, Boehnke M, Abecasis GR. Low-coverage sequencing: implications for design of complex trait association studies. Genome Res. 2011; 21: 940-951.

37. Pasaniuc B, Rohland N, McLaren PJ, Garimella K, Zaitlen N, Li H, et al. Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nat Genet. 2012; 44: 631-635.

38. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet. 2010; 11: 773-785.

39. Asimit J, Zeggini E. Rare variant association analysis methods for complex traits. Annu Rev Genet. 2010; 44: 293-308.

40. Consortium TIH. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005; 437: 1299–1320.

41. Asimit JL, Day-Williams AG, Morris AP, Zeggini E. ARIEL and AMELIA: testing for an accumulation of rare variants using next-generation sequencing data. Hum Hered. 2012; 73: 84-94.

42. Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). Mutat Res. 2007; 615: 28–56.

43. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet. 2008; 83: 311-321.

44. Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genet Epidemiol. 2010; 34: 188-193.

45. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009; 5: e1000384.

46. Zawistowski M, Gopalakrishnan S, Ding J, Li Y, Grimm S, Zollner S. Extending rare-variant testing strategies: analysis of noncoding sequence and imputed genotypes. Am J Hum Genet. 2010; 87: 604-617.

47. Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, et al. Testing for an unusual distribution of rare variants. PLoS Genet. 2011; 7: e1001322.

48. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics. 2012; 13: 762-775.

49. Basu S, Pan W. Comparison of statistical tests for disease association with rare variants. Genet Epidemiol. 2011; 35: 606-619.

50. Han F, Pan W. A data-adaptive sum test for disease association with multiple common or rare variants. Hum Hered. 2010; 70: 42-54.

51. Hoffmann TJ, Marini NJ, Witte JS. Comprehensive approach to analyzing rare genetic variants. PLoS One. 2010; 5: e13584.

52. Lin DY, Tang ZZ. A general framework for detecting disease associations with rare variants in sequencing studies. Am J Hum Genet. 2011; 89: 354-367.

53. Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, et al. Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet. 2010; 86: 832-838.

54. Liu DJ, Leal SM. A novel adaptive method for the analysis of next-generation sequencing data to detect complex trait associations with rare variants due to gene main effects and interactions. PLoS Genet. 2010; 6: e1001156.

55. Ionita-Laza I, Buxbaum JD, Laird NM, Lange C. A new testing strategy to identify rare variants with either risk or protective effect on disease. PLoS Genet. 2011; 7: e1001289.

56. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011; 89: 82-93.

57. Pan W. Asymptotic tests of association with multiple SNPs in linkage disequilibrium. Genet Epidemiol. 2009; 33: 497-507.

58. Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, et al. Testing for an unusual distribution of rare variants. PLoS Genet. 2011; 7: e1001322.

59. Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, et al. Powerful SNP-set analysis for case-control genome-wide association studies. Am J Hum Genet. 2010; 86: 929-942.

60. Davies RB. Algorithm AS 155: The distribution of alinear combination of c 2 random variables. J R Stat Soc Ser C Appl Stat. 1980; 29: 323–333.

61. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. NHLBI GO Exome Sequencing Project—ESP LungProject Team. Optimal unified approach for rare variant association testing with application to small-samplecase-control whole-exome sequencing studies. Am J Hum Genet. 2012; 91: 224–237.

62. Derkach A, Lawless JF, Sun L. Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests. Genet Epidemiol. 2013; 37: 110-121.

63. Fisher RA, Genetiker S. Statistical methods for research workers. 1970.

64. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics. 2012; 13: 762-775.

65. Chen LS, Hsu L, Gamazon ER, Cox NJ, Nicolae DL. An exponential combination procedure for set-based association tests in sequencing studies. Am J Hum Genet. 2012; 91: 977-986.

66. Dering C, Konig IR, Ramsey LB, Relling MV, Yang W, Ziegler A. A comprehensive evaluation of collapsing methods using simulated and real data: excellent annotation of functionality and large sample sizes required. Frontiers in Genetics. 2014; 5: 323.

67. Almasy L, Dyer TD, Peralta JM, Kent JW, Charlesworth JC, Curran JE, et al. Genetic Analysis Workshop 17 mini-exome simulation. BMC Proc. 2011; 5: 2.

68. Trevino LR, Shimasaki N, Yang W, Panetta JC, Cheng C, Pei D, et al.

Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. J Clin Oncol. 2009; 27: 5972-5978.

69. Bhatia G, Bansal V, Harismendy O, Schork NJ, Topol EJ, Frazer K, et al. A covering method for detecting genetic associations between rare variants and common phenotypes. PLoS Comput Biol. 2010; 6: e1000954.

70. Luo L, Boerwinkle E, Xiong M. Association studies for next-generation sequencing. Genome Res. 2011; 21: 1099-1108.

71. Zhang Q, Irvin MR, Arnett DK, Province MA, Borecki I. A data-driven method for identifying rare variants with heterogeneous trait effects. Genet. Epidemiol. 2011; 35: 679–685.

72. Derkach A, Lawless JF, Sun L. Pooled association tests for rare genetic variants: a review and some new results. Stat Sci. 2014; 29: 302–321.

Austin Biom and Biostat - Volume 2 Issue 1 - 2015
Submit your Manuscript | www.austinpublishinggroup.com

Citation: Padhukasahasram B. Region-Based Tests for Association Analysis of Rare Variants. Austin Biom and Biostat. 2015;2(1): 1012.