|  |
| --- |
| modSaRa2: An Accurate and Powerful Method for Copy Number Variation Detection |
| Integration of multiple genetic sources for copy number variation detection is a powerful approach to improve the identification of variants associated with complex traits. Although it has been shown that the widely used change-point based methods can increase statistical power to identify variants, it remains challenging to effectively identify CNVs with weak signals due to the noisy nature of genotyping intensity data. We previously developed modSaRa, a normal mean-based model on a screening and ranking algorithm for copy number variation identification which presented desirable sensitivity with high computational efficiency. Here we proposed a novel improvement by integrating the relative allelic intensity with prior information of statistics into modeling to boost statistical power for the identification of variants, so called modSaRa2. Simulation studies illustrated that modSaRa2 markedly improved both sensitivity and specificity over existing methods. The improvement for weak CNV signals is the most substantial, while simultaneously improving stability when CNV size varies. The application of the new method to a whole genome melanoma dataset identified novel candidate melanoma risk associated CNV variants which facilitates understanding of the possible roles of germline copy number variants in development of melanoma. |