Hill-climbing through "random chemistry" for detecting epistasis

Margaret J. Eppstein
Dept. of Computer Science
University of Vermont
Burlington, VT 05405
802-656-1918
Maggie.Eppstein@uvm.edu

Joshua L. Payne
Dept. of Computer Science
University of Vermont
Burlington, VT 05405
802-656-9116
Joshua.Payne@uvm.edu

Bill C. White
Computational Genetics Laboratory
Dartmouth College
Lebanon, NH 03756
Bill.C.White@dartmouth.edu

Jason H. Moore
Computational Genetics Laboratory
Dartmouth College
Lebanon, NH 03756
603-653-9939
Jason.H.Moore@dartmouth.edu

ABSTRACT
There are estimated to be on the order of 10^6 single nucleotide polymorphisms (SNPs) existing as standing variation in the human genome. Certain combinations of these SNPs can interact in complex ways to predispose individuals for a variety of common diseases, even though individual SNPs may have no ill effects. Detecting these epistatic combinations is a computationally daunting task. Trying to use individual or growing subsets of SNPs as building blocks for detection of larger combinations of purely epistatic SNPs (e.g., via genetic algorithms or genetic programming) is no better than random search, since there is no predictive power in subsets of the correct set of epistatically interacting SNPs. Here, we explore the potential for hill-climbing from the other direction; that is, from large sets of candidate SNPs to smaller ones. This approach was inspired by Kauffman's "random chemistry" approach to detecting small autocatalytic sets of molecules from within large sets. Although the algorithm is conceptually straightforward, its success hinges upon the creation of a fitness function able to discriminate large sets that contain subsets of interacting SNPs from those that don't. Here, we employ an approximate and noisy fitness function based on the ReliefF data mining algorithm. Preliminary results from synthetic data sets show that the resulting algorithm can detect epistatic pairs from up to 1000 candidate SNPs in O(log *N*) fitness evaluations, although success rate degrades as heritability declines. The results presented herein are offered as proof of concept for the random chemistry approach, but research continues into seeking a more accurate fitness approximator for large sets that will enable us to extend the approach to larger data sets and to lower heritabilities.

GECCO'06, July 8–12, 2004, Seattle, WA, USA.
ACM 1-59593-186-4/06/0007.

Categories and Subject Descriptors
I.2.8 [**Computing Methodologies**]: Problem Solving, Control Methods, and Search

General Terms
Algorithms, Design.

Keywords
Population based optimization, epistasis, SNPs, data mining.

1. INTRODUCTION
There are estimated to be on the order of 10^6 single nucleotide polymorphisms (SNPs) existing as standing variation in the human genome, and current genetic studies are identifying many of these. While the majority of these SNPs have no physiological effect, certain combinations of these SNPs can interact in complex ways to predispose individuals for a variety of common diseases, even though the individual SNPs may have no ill effects. Increasingly, researchers are seeking new methods that will enable us to capitalize on the rapidly growing databases of human variability by conducting genome-wide association studies that could revolutionize detection, prevention, and treatment of many common complex diseases [1],[2]. However, detecting which handfuls of SNPs (from among hundreds, thousands, or even hundreds of thousands of candidate SNPs) exhibit nonlinear epistatic interactions is a computationally daunting optimization task [3]. Solution of this important problem is further exacerbated by low disease heritability, small sample sizes, little or no marginal effects for individual SNPs, and unknown *a priori* information regarding how many, if any, SNPs interact epistatically. An analytical retooling is needed to deal with these challenges [4] and some new methods, such as multifactor dimensionality reduction or MDR [5]-[7], show promise [8]. Despite some progress in this area, the optimal search or wrapper method for implementing these methods for genome-wide genetic analysis is still an open question. Exhaustive searches are not

computationally feasible, gradient-based algorithms are not suitable, and stochastic search algorithms will require careful design to fully embrace the difficult nature of the fitness landscape [9]-[10].

In this work, we propose a new approach, inspired by Stuart Kauffman's "random chemistry" [12], that tackles the problem by hill-climbing from larger to smaller sets of SNPs, rather than vice versa. First, we discuss how to evaluate the predictive power of a set of SNPs for a particular disease, and briefly review why standard evolutionary algorithms are ill-suited to locating epistatic combinations of SNPs. We then describe the proposed random chemistry approach, which necessitates the creation of a different type of fitness metric. Preliminary tests using synthetic data sets with 0.4 and 0.1 heritabilities establish proof-of-concept for the method, and future directions are discussed.

2. EVALUATING FITNESS

How does one evaluate whether a small set of SNP loci interact to influence susceptibility to a given condition? This is a non-trivial question, especially when heritability is low and epistatic interactions are such that different genotypes at the same loci exhibit different penetrance values for the same disease. (Without loss of generality, we refer to the condition in question generically as the "disease".) Here, we borrow an idea commonly employed in medical decision making; i.e., the receiver operating characteristic (ROC) curve [11], described below. The "sensitivity" of a test is equivalent to the true positive fraction (TPF); that is, the fraction of test subjects which have the disease and yield a positive test result. On the other hand, the "specificity" of a test is simply 1 minus the false positive fraction (FPF), where the FPF is the fraction of test subjects that do not have the disease but still test "positive". An ROC curve is simply a plot showing the trade-off between sensitivity and specificity (typically plotted as 1-specificity) as we vary some cutoff criterion for when a test is considered "positive" (e.g., Figure 1). Ideally, a good test will have high sensitivity and high specificity,

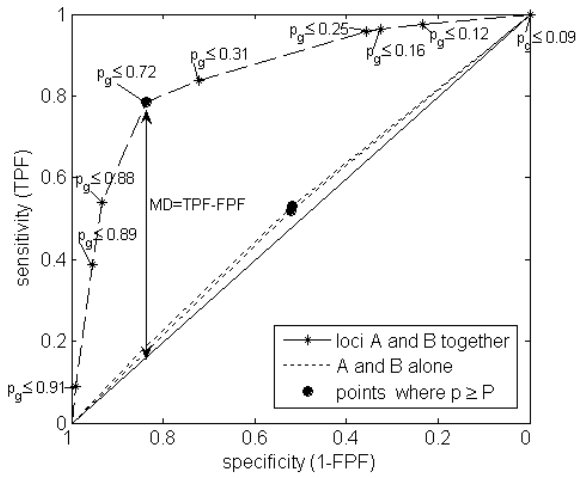

Figure 1. ROC curve for the data shown in Table 1, with the maximum distance (MD) fitness metric shown for the two interacting A and B loci. Note that the specificity axis is shown in decreasing order.

so will have an ROC curve that passes close to the upper left hand corner of the plot. The area under the curve (AUC) is frequently used as a measure of the predictive power of a test. A test with no predictive power will simply have a diagonal ROC curve, with AUC = 0.5.

In the case of assessing the predictive power of a set of SNPs, we can create an ROC curve as follows. First, we compute the "sample penetrance" p_g for each possible diploid genotype g at the specified SNP loci; i.e., the proportion of subjects in the sample with genotype g that exhibit the disease. The ROC curve can then be estimated by varying the cutoff of sample penetrance that is considered a "positive" test. Since there are only a finite number of possible genotypes for a given set of loci, this is not a continuous curve, but rather is a set of discrete points. Thus, it is not meaningful to try to calculate an AUC. Instead, we use the maximum distance (*MD*) above the diagonal as a measure of the predictive power of a set of SNPs. The "sample prevalence" P is the proportion of the sample that exhibits the disease. The *MD* occurs at the point where the cutoff includes all genotypes whose sample penetrance is greater than or equal to the sample prevalence, as follows:

$$MD = TPF - FPF, \forall g \left| p_g \ge P \right. \quad (2.1)$$

We illustrate the ROC curve and the *MD* fitness metric with a synthetic data set that contains 1000 SNPS, each with two possible allele values, from 1600 individuals, half of whom have disease (so P=0.50). Although none of the SNP loci are individually correlated with the disease, the alleles at two loci interact epistatically to affect susceptibility to the disease with 0.4 heritability. Suppose the two epistatically interacting SNP loci are locus A, with 3 possible diploid genotypes *AA*, *Aa*, and *aa*, and locus B, with genotypes *BB*, *Bb*, and *bb*. Then there are 3^2 = 9 possible genotype combinations. In this example, the calculated sample penetrance for these 9 genotypes is as shown inside the double lines in Table 1, with the boldface values representing those genotypes with sample penetrance ≥ sample prevalence. Note that sample penetrance values for the A and B loci taken individually (right column and bottom row, respectively) are all close to the sample prevalence of 0.5 (bottom right value), indicating that there are no marginal effects.

Table 1. Sample penetrance p_g for all 9 genotypes of two loci, A and B, that interact epistatically to influence susceptibility to a disease with heritability 0.4.

p_g	*BB*	*Bb*	*bb*	*A alone*
AA	0.09	**0.72**	**0.91**	0.52
Aa	**0.89**	0.25	0.31	0.48
aa	0.12	**0.88**	0.16	0.53
B alone	0.48	0.52	0.48	P=0.50

The ROC curve for this data is shown in Figure 1, showing how the points on the curve correspond to the values in the table, and depicting the *MD* fitness metric for the epistatic loci. Also shown, for comparison, are the ROC curves for the A and B loci alone, which are not significantly above the main diagonal and

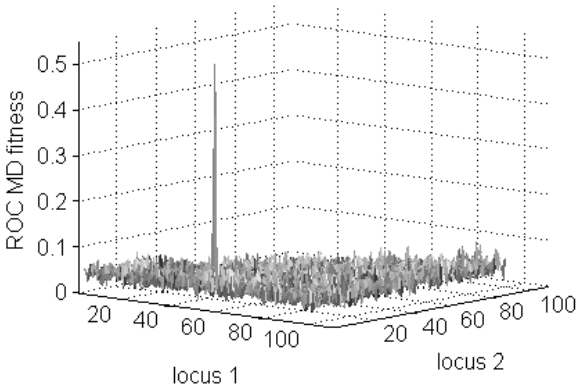

Figure 2. A portion of the needle-in-a-haystack fitness landscape evaluating the predictive power of all possible pairs in SNPs; only the correct pair has high *MD* fitness, and thus hill-climbing towards the correct pair is not possible.

hence indicate that these loci have no predictive power when viewed in isolation. Assuming that there are at most two allelic variants at a given locus (which is typically the case), then there are 3 possible diploid genotypes per locus (the two homozygotes and the heterozygote), so there are 3^L possible genotype combinations at L loci for which the sample penetrance must be computed. Thus, computing the *MD* fitness metric for L loci has time complexity $\theta(3^L)$, which is only practical for small L.

If individual loci have no significant marginal effects (as in the synthetic data sets used herein), then it is not possible to hill-climb from smaller sets of SNPs with no detectable genetic effects to larger sets of SNPs that exhibit epistatic interactions. Thus, evolutionary algorithms such as genetic algorithms or genetic programming, that attempt to grow small, high fitness building blocks into complete solutions, will be no better than random search [9],[10]. To illustrate this point, Figure 2 shows a portion of this needle-in-a-haystack fitness landscape, for a variety of pairs of loci, only two of which interact.

If hill-climbing from lower to higher order building blocks is not possible, then are we stuck with brute force or random search? In order to try all combinations of $2..L_{max}$ out of N possible loci, brute force algorithms will require $\theta\left(\sum_{L=2}^{L_{max}}\binom{N}{L}\right)$ fitness evaluations, each of which requires $\theta(3^L)$ time to compute. Clearly, this becomes computationally prohibitive for large N and/or L. In this work, we explore an alternative approach that attempts to hill climb from larger sets of SNPs to smaller sets of SNPs using a proposed algorithm we dub "random chemistry", described in the following sections.

3. RANDOM CHEMISTRY ALGORITHM

Stuart Kauffman [12] outlined a simple procedure for detecting small autocatalytic sets of molecules from among large numbers of potential molecules. The process he described, to search for L interacting molecules from N potential molecules, can be summarized as shown in Algorithm 1.

Note that, if we grab a random 50% of the molecules, the probability that we got any given molecule is 0.5. Thus, the probability that we got lucky and got all L of the desired molecules is 0.5^L, and hence we would expect that, on average, one out of every 2^L tubes will contain all of the L molecules that we seek. By using $\sigma \cdot 2^L$ tubes, where σ is a safety margin to help ensure that at least one tube has all L molecules (Kauffman suggested σ=2), we increase the probability that at least one tube will be "positive". Note that the algorithm will require only $\theta(\log N)$ screening tests (i.e., fitness evaluations) to pick out the L interacting molecules.

Algorithm 1: "Random Chemistry"

1) Put a random 50% of the "parent set" of molecules into each of $\sigma \cdot 2^L$ test tubes (the "child" sets).
2) Determine which of the child sets are "positive" (i.e., contain the autocatalytic set of L molecules) by screening each of the tubes for some resulting product of the catalytic reaction.
3) Save only those molecules that are in one of the "positive" tubes (i.e., replace the parent set with a positive child set).
4) Repeat from Step 1, until the number of molecules has been reduced to L.

Can we adopt a "random chemistry" approach to detecting small sets of L epistatically interacting SNPs from among N candidate SNPs? The answer depends on whether or not we can effectively do the "screening" process in Step 2. In other words, can we construct a fitness function that yields higher fitness for large sets of SNPs that contain the correct subset of L SNPs from those that don't? Furthermore, if we do not know exactly the degree of epistasis L, then can our fitness function yield higher fitness for smaller sets of SNPs that contain the correct subset of L SNPs than for larger sets that contain the correct subset of L SNPs, so that we can simultaneously identify both the number and location of epistatically interacting SNPs? These questions are explored in the next section.

4. FITNESS REVISTED

In Section 2 we described the *MD* fitness metric. Is it sufficient for use in a random chemistry algorithm? Unfortunately, the answer is "no". Although *MD* is higher for sets containing the correct subsets of epistatically interacting SNPs, it continues to increase as additional "extra" SNPs are considered. This is illustrated in Figure 3a for the same synthetic 1600 member data set described in Section 2, for random sets containing both correct loci (×'s) and random sets containing 0 or 1 of the correct loci (dots). The reason for this is two-fold; first, adding in additional loci does not lower the predictive power of epistatically interacting loci also contained in the set, and second, as the number of loci L in a set increases, the number of genotypes increases as 3^L. As the number of possible genotypes approaches the number of samples in the data set, we are simply over-fitting the data and so the *MD* fitness metric continues to increase. In order to compensate for overfitting, we incorporate cross-validation into the fitness metric. Specifically, we divide the sample in two, and compute sample penetrance tables for each half of the data set. For each of these two tables, we determine which of the 3^L genotypes a) were represented in both tables, b) had sample penentrance < the sample prevalence ("negatives"),

c) had sample penentrance ≥ the sample prevalence ("positives"). We then define a cross-correlation value C (Figure 3b) to be the proportion of genotypes that "agreed" between the two tables (i.e., were either both positively or both negatively associated with the disease), and we also define a metric of support S (Figure 3c) as the proportion of the 3^L genotypes that had representatives in both tables. We define m to be the smallest MD above the ROC diagonal that is considered to be more predictive than random (we used $m = 0.1$). If $MD > m$, then we repeat the cross-validation test for some number of repetitions (we used 20). More sophisticated types of cross-validation could certainly be employed, if necessary. A new fitness function F is empirically defined from the averaged results, as follows:

$$F = \left(\overline{MD} - m\right) \cdot \overline{C} \cdot \overline{S}^{0.25} \qquad (2.2)$$

The support S is raised to the 0.25 power to minimize the effects of the strong non-linearity in this metric (Figure 3c). While none of MD, C, or S alone satisfy the criteria for our fitness function (Figure 3a,b,c), the combined fitness metric F does, at least for relatively small L (Figure 3d), enabling us to distinguish larger sets of SNPs that contain the correct loci from those that don't, and enabling us to hill-climb from larger sets to smaller sets that contain fewer "extra" loci.

Unfortunately, the fitness function defined in equation (2.2) is only useful on these data sets for $L \leq 8$, since a) with sample size of 1600, the support is already close to zero for $L = 8$ (Figure 3c), and b) because the fitness computation is $\theta(3^L)$ equation (2.2) becomes too slow to be practical for $L > 8$. However, for the random chemistry algorithm to be useful we will need to evaluate the fitness of sets of SNPs in the hundreds or thousands. This remains a nontrivial problem and we are continuing to seek better ways to estimate the fitness of large sets of SNPs. However, to date, the most effective fitness approximator we have achieved is based on the ReliefF data mining algorithm [13], which has previously been shown to have promise for the SNP problem [14]. The ReliefF algorithm attempts to estimate importance of weights of each locus in discriminating between two classes (e.g., healthy and diseased) as shown in Algorithm 2.

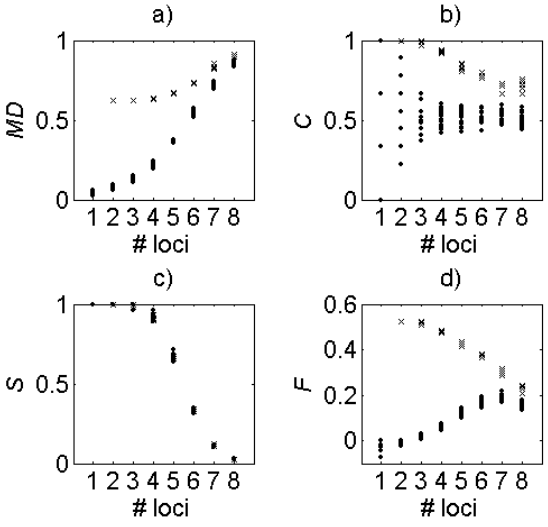

Figure 3. a) Maximum distance metric, b) cross validation metric, c) support metric, and d) fitness metric per equation (2.2) for the synthetic data set described in Section 2. Random subsets of SNPs containing both correct loci are shown with ×'s and random sets containing 0 or 1 of the correct loci are shown with dots.

Algorithm 2: "ReliefF"

Weights W_j=0, $\forall\, j \in \{1..N\}$ loci

For i = 1 to k (k = # random samples)

  Select an individual R_i

  Hits = *nn* nearest neighbors from same class as R_i

  Misses = *nn* nearest neighbors from other class as R_i

  For j = 1 to N (for all loci)

    H_j = prop. of the *nn Hits*$_j$ that matched value of R_{ij}

    M_j = prop. of the *nn Misses*$_j$ that matched value of R_{ij}

    $W_j = W_j + (H_j - M_j)/k$ (estimate importance of each locus)

  End for

End for

We define a "rough" fitness function F_r as:

$$F_r = \text{ mean of top 25\% of Weights } W \qquad (2.3)$$

using the weights W as calculated by ReliefF (Algorithm 2). For the results reported here, we used k = 1600 (where we used each sample individual exactly once) and nn = 10 (i.e., 10 nearest neighbors). The nearest neighbors are determined by maximizing

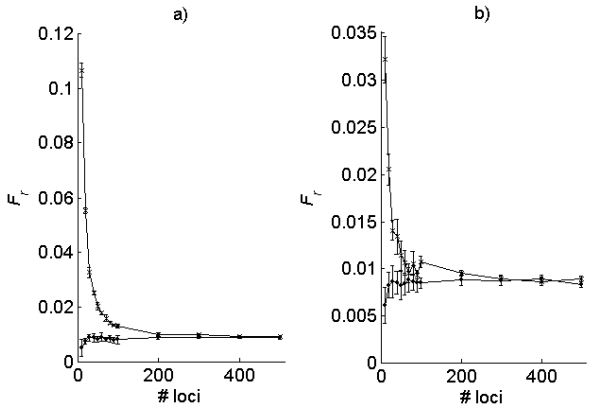

Figure 4. The rough fitness approximator F_r from equation (2.3) as applied to synthetic data with heritability a) 0.4 and b) 0.1. The top lines are for random sets that contain the correct 2 loci and the bottom lines are for random sets without the correct loci. Each data point is the mean of 5 random sets, with error bars indicating ± one standard deviation.

the number of loci with the same genotype as the sample R_i. For large numbers of loci N this fitness approximator gets increasingly noisy, because without knowing which loci are the most important SNPs, the ReliefF algorithm may pick the "wrong" nearest neighbors (i.e., based on matching genotypes of irrelevant loci). Nonetheless, this fitness approximator works fairly well up to a few hundred SNPs, although it becomes noisier as heritability decreases (Figure 4).

5. COMPENSATING FOR NOISE

Since F_r is a noisy approximate fitness function, we modified Step 3 of the random chemistry algorithm to save, not just the single fittest child set, but the top t fittest children sets, where $t = \lfloor \log_4 N_p \rfloor$ and N_p is the number of SNPs in the parent set of the current iteration. These sets were then merged by saving all of the SNPs in the fittest set along with all other SNPs that occurred in at least two of the remaining top sets. Thus, set sizes are no longer reduced by exactly ½ during each iteration, but were only reduced by approximately ½ to ¼ (Figure 5). We continued iterating the random chemistry algorithm until the set size was reduced to only 8 SNPS, after which we applied a brute force evaluation of all possible combinations of 2 through 8 SNPs, using the more accurate fitness metric F from equation (2.2). In practice, this requires on the order of $\log_{1.5}$N iterations (rather than $\log_2$N iterations, as in algorithm 1), due to the merging of sets.

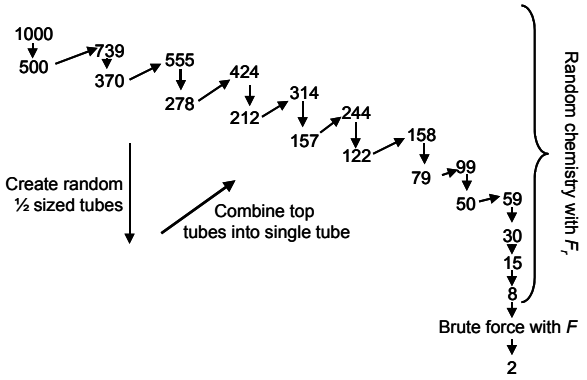

Figure 5. Decrease in SNP set sizes during a representative run of the random chemistry algorithm compensated for noise, in which 2 epistatically interacting SNPs were correctly detected from 1000 SNP loci, using the data set described in Section 2.

6. EXPERIMENTAL RESULTS

We have performed preliminary testing of the algorithm using synthetic data generated as described in [15]. In each data set, there were 2 epistatically interacting SNPs with no marginal effects. There were 1600 samples in each set with sample prevalence of 0.5, and major/minor alleles frequencies of 0.6/0.4, respectively, at each locus. Some of the data sets had disease heritability of 0.4, and some had heritability of 0.1. For all data sets, we created 12 child sets from the parent set, at step 1 of the random chemistry algorithm (Algorithm 1). At total of 250 test runs were performed, as follows. We ran 5 repetitions of 5 different data sets with each of 200, 500, and 1000 SNPs, for a total of 75 test runs at 0.4 heritability. In addition, we ran 5 repetitions of 5 different data sets with each of 100, 150, 200, 300, 400, 500, and 1000 SNPs, for a total of 175 test runs at 0.1 heritability. Since F_r is so much noisier for the 0.1 heritability data than at 0.4 heritability (Figure 4) we ran more tests at the lower heritability.

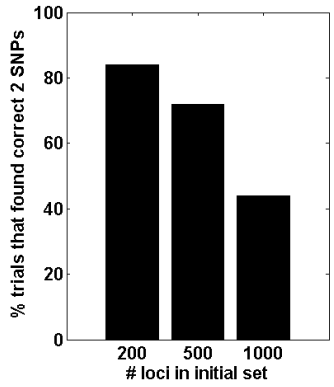

Figure 6. Percent of the time that 2 epistatically interacting SNPs, with disease heritability 0.4, were correctly isolated from initial sets of varying sizes. Each bar represents 5 trials on each of 5 different data sets. For each initial set size, at least one of the 5 trials on a given data set was able to find the correct 2 SNPs. Note that the x-axis is not uniformly spaced.

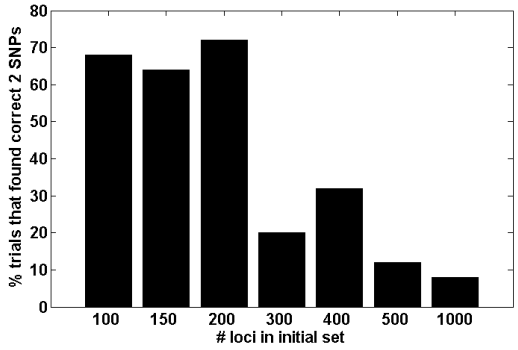

Figure 7. Percent of the time that 2 epistatically interacting SNPs, with disease heritability 0.1, were correctly isolated from initial sets of varying sizes. Each bar represents 5 trials on each of 5 different data sets. For initial set sizes up to 200, at least one of the 5 trials on a given data set was able to find the correct 2 SNPs. Note that the x-axis is not uniformly spaced.

For the 0.4 heritability data, the percent of trials that were able to correctly isolate the 2 epistatically interacting SNPs declined with the initial set size (Figure 6) but, even with 1000 initial SNPs, at least one of the the five trials on a given data set was able to correctly identify the 2 SNPs. Thus, doing repeated trials can further compensate for noise in the fitness function. For the 0.1 heritability data, the percent of trials that were able to correctly isolate the 2 epistatically interacting SNPs dropped more rapidly with increasing initial set size (Figure 7). For the lower

heritability data, at least one of the five trials on a given data set was able to correctly identify the 2 SNPs from initial set sizes up to 200 SNPs, but for larger set sizes there were some data sets for which none of the 5 trials found the correct 2 SNPs. Overall, these results are better than we would have predicted, given the level of noise in the rough fitness approximator (Figure 4).

7. SUMMARY AND FUTURE WORK

Detecting small sets of epistatically interacting SNPs that can influence an individual's susceptibility to disease is a difficult but important challenge facing bioinformaticists. Since smaller subsets of purely epistatically interacting sets of SNPs have no effect, it is not possible to hill-climb from smaller to larger building blocks in constructing these SNP sets. In this work we have presented an alternative population based approach dubbed the "random chemistry" algorithm, which enables us to hill-climb from larger to smaller sets of SNPs in only O(log N) fitness evaluations, where N is the number of SNPs in the initial set. Furthermore, the algorithm is inherently parallelizeable. The main challenge in implementing this algorithm is in finding a reliable way to approximate which large sets contain the correct SNPs and which don't. Herein we employ an approximate and noisy fitness function based on the ReliefF data mining algorithm. Although this fitness approximator is far from ideal, using it in combination with noise compensation techniques has enabled us to pick out 2 epistatically interacting SNP loci from up to 1000 candidate loci, although not surprisingly the success rate declines with declining heritability.

The results presented herein are offered as proof of concept for the random chemistry approach, but research continues into seeking a more accurate fitness approximator for large sets that will enable us to extend the approach to larger data sets and to lower heritabilities. In particular, we are investigating the use of artificial neural networks or support vector machines as more accurate rough fitness approximators. In real data sets, there may also be additional information which can be used to pre-filter the set of candidate SNPs, such as the presence of small marginal effects of individual loci, proximity of SNPs to coding or regulatory regions of genes suspected to be associated with the disease, etc. These pre-filtering approaches can help to reduce initial set sizes down to a size that can be tackled through random chemistry. We plan to apply the algorithm to actual SNP data sets in the coming months. The random chemistry algorithm may also prove useful for related search problems in other domains, such as in microarray analysis.

ACKNOWLEDGMENTS

This work was supported, in part, by a pilot award and a graduate research assistantship, funded by DOE-FG02-00ER45828 awarded by the US Department of Energy through its EPSCoR Program and by National Institutes of Health grants AI59694 and LM009012. We thank Joshua Gilbert for his aid in creating the synthetic data sets.

8. REFERENCES

[1] Hirschhorn, J.N, and Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* **6**, 95-108 (2005)

[2] Wang, W.Y., Barratt, B.J., Clayton, D.G., Todd, J.A. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*. **6**(2):109-18 (2005)

[3] Moore, J.H., Ritchie, M.D. The challenges of whole-genome approaches to common diseases. *JAMA* **291**, 1642-1643 (2002).

[4] Thornton-Wells TA, Moore JH, Haines JL. Genetics, statistics and human disease: analytical retooling for complexity. *Trends Genet.*, **20**(12):640-7. (2004)

[5] Ritchie, M.D., Hahn, L.W., Roodi, N., Bailey, L.R., Dupont, W.D., Parl, F.F., and Moore, J.H. Multifactor dimensionality reduction reveals high-order interactions among estrogen metabolism genes in sporadic breast cancer. *Amer. J. of Hum. Gen*. **69**, 138-147. (2001)

[6] Moore, J.H. Computational analysis of gene-gene interactions in common human diseases using multifactor dimensionality reduction. *Expert. Rev. Mol. Diagn.* **4**, 795-803 (2004).

[7] Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, White BC. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J Theor Biol.*, in press. (2006)

[8] McKinney BA, Reif DM, Ritchie MD, Moore JH. Machine learning for detecting gene-gene interactions. *Applied Bioinformatics*, in press. (2006)

[9] White, B.C., Gilbert, J.C, Reif, D.M., Moore, J.H. A statistical comparison of grammatical evolution strategies in the domain of human genetics. *Proc. of the IEEE Congress on Evol. Computing*, 676-682 (2005).

[10] Moore, J.H. and White, B.C. Genome-wide genetic analysis using genetic programming: The critical need for expert knowledge. In *Genetic Programming Theory and Practice IV, Springer*, in press (2006).

[11] Barrett, H.H, and Myers, K.J. *Foundations of Image Science,* John Wiley & Sons, Inc., NJ (2004).

[12] Kauffman, S. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*, Oxford Univ. Press, USA, 336 pp. (1996).

[13] Robnik-Sikonja, M., and Konenenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learning*. **53**, 23-69 (2003).

[14] Moore, J.H. and White, B.C.. Tuning ReliefF for genome-wide genetic analysis. *Submitted*. (2006)

[15] Culverhouse, R., Suarez, B.K., Lin, J., and Reich, T. A perspective on epistasis: Limits of models displaying no main effect. *Am. J. Hum. Genet*. **70**, 461-471 (2002).