Structured Populations with Limited Resources Exhibit Higher Rates of Complex Function Evolution

Arthur W Covert III[1,2], Siena McFetridge[1,2] and Evan DeLord[1,2]

[1]University of Texas at Austin, Center for Computational Biology and Bioinfomatics
[2]Beacon Center for the Study of Evolution in Action
art.covert@austin.utexas.edu

Abstract

The impact of population structure on evolving populations is difficult to study. Populations broken up into groups of organisms and connected by low levels of migration will experience different types of geneflow than normal unstructured populations. Various studies, spanning decades of research, have lead to seemingly contradictory conclusions. Some point to population structure as a means to improve adaptation, others argue that population structure hinders evolution. We investigate how population structure impacts the evolution of complex functions in environments with limited resources. We find that structured populations with limited resources tend to evolve complex functions at a higher rate than unstructured populations, across a broad range of migration rates. This suggests that population structure may have an important impact on evolution, in both sexual and asexual populations, at least at certain migration rates.

Introduction

Few topics in evolutionary biology have been as hotly debated as the impact of population structure (Moore and Tonsor 1994, Phillips 1996, Coyne et al. 1997, Wade and Goodnight 1998, Coyne et al. 2000, Goodnight and Wade 2000, Kryazhimskiy et al. 2012, Covert and Wilke 2014). Sewall Wright first suggested that populations structured into isolated subpopulations, connected with low levels of migration, created a balance between exploration of fitness landscapes and exploitation of fitness peaks (Wright 1932). According to Wright's theory, these structured populations could pass through fitness valleys that large unstructured populations could not.

Evolutionary computation researchers have relied on structured populations to improve the speed of evolutionary search, though little work has been done to determine if the improvement comes from a process similar to the one Wright envisioned, or some other source (Lin et al 1994, Belding 1995, Cantu-Paz 1998, Cantu-Paz 2001, Fernándex et al 2003). Meanwhile, the notion that structured populations may improve adaptation has proved controversial in evolutionary biology (Conye et al 1997, Coyne et al 2000), with many scientist arguing that populations structure may have no effect, or a negative effect on rates of adaptation (Kryazhimskiy 2012). Other studies have suggested that population structure may have a very important impact under certain conditions (Moore and Tonsor 1994, Wade and Goodnight 1998, Goodnight and Wade 2000, Covert and Whilke 2014).

In structured populations, beneficial mutations take longer to sweep because they must wait for migration to carry them to every subpopultion. The longer a beneficial mutation takes to sweep, the greater the chance that a superior beneficial mutation will be discovered in one of the subpopulations that has not yet been swept yet. This process has been observed in large structured asexual populations and is referred to as leapfrogging (Gerrish and Lenski 1998, Miller et al 2011, Covert and Wilke 2014).

In a recent paper, Covert and Wilke (2014) demonstrated that at certain migration rates, leapfrogging from older genetic backgrounds yielded a dramatic fitness improvement. However, these experiments were asexual and did not address the impact of population structure in sexual populations. In addition, Covert and Wilke's experiments simulated populations with unlimited resources, which was a necessary simplifying assumption for their work, but is ultimately biologically unrealistic.

In this paper, we examine the impact of population structure in environments with limited resources. We run experiments with self-replicating digital organism that may reproduce asexually or sexually. We demonstrate that population structure improves the adaptive ability of populations, even when resources are limited, under both methods of reproduction.

Methods

We used Avida 2.12.2 (Ofria and Wilke 2004), to evolve self-replicating digital organisms, our complete experimental setup is free to download from git hub. Digital organisms are computer programs that exist on a grid. Each grid cell contains a virtual CPU and memory for the organisms to execute their code. During their lifetime each organism must copy every instruction in its genome into a new memory space and then execute a special "divide" command that places the new offspring in the environment. Single point mutations had a 25% probability of occurring on divide. All of our

populations were seeded with a hand-written ancestor 51 instructions long that could do nothing initially but copy itself.

All environments had nine logical functions that were rewarded in terms of increasing complexity and resource availability (Lenski et al 1999, Chow et al 2001, Covert et all 2012). Digital organisms could perform each of the nine logical functions by evolving a series of NAND operations. When a function was successfully performed the organisms' were rewarded only if there was sufficient resource in their current cell. When sufficient resources were present, they would be consumed by the execution of the function and the organism would be allowed to run its genome faster (Table 1). Rewards for each logical function were proportional to the number of NAND operations required to execute the function's most parsimonious configuration. Standard Avida populations have a difficult time evolving the most complex functions: XOR and EQU (Lenski et al. 2003, Covert et al 2012).

Function Name	Logic Operation	Speed Increase
NOT	~A; ~B	x2
NAND	~(A AND B)	x2
AND	A AND B	x4
OR_N	(A OR ~B) (~A OR B)	x4
OR	A OR B	x8
AND_N	(A AND ~B) (~A AND B)	x8
NOR	~A AND ~B	x16
XOR	(A AND ~B) OR (~A AND B)	x16
EQU	(A AND B) OR (~A AND ~B)	x32

Table 1 The standard nine logical functions in the Avida environment and their speed increases. Digital organisms have only the NAND operation in their instruction sets and must construct other logical functions out of NAND operations. The energy bonus for each function is equivalent to 2^n, where n is the minimum number of NAND operations needed to complete it. In our experiments, each function reward was granted only if the organism was able to consume 1 unit of a limited resource.

We ran two experiments to test the effects structured populations with limited resources, one sexual experiment and one asexual experiment. Each experiment had 4 treatments, 3 in which the population was structured into 100 demes each with 100 organisms, and 1 that was a single large deme of 10,000 organisms. Each treatment had 40 replicates, seeded with the default ancestor and a unique random number seed.

Migration between demes occurred when digital organisms were born. Each population ran for 250,000 updates.[1]

Each organism had a small probability of being placed in a different deme when it divided from its parent. Migration rates were selected based on previous results (Covert and Wilke 2014) indicating that the optimal migration rate was around $5x10^{-5}$ (one migration every other generation). We ran additional migration rates at one order of magnitude higher, $5x10^{-4}$ (5 migrants every generation) and one order of magnitude lower, $5x10^{-6}$ (one migrant every 20 generations).

In the first experiment, all organisms reproduced asexually, with offspring being placed in an adjacent grid cell. In the second experiment, all organisms reproduced via sexual recombination, similar to the recombination in Misivic et al (2006). In the sexual experiments, organisms copied their genomes and then placed their offspring into a "birth chamber". Each birth chamber was associated with the organism's deme. When an organism was placed in the birth chamber, it waited to be joined by a second offspring organism. When two organisms were present in the birth chamber, their genetic code was divided up into 5 regular segments, each segment had a 50% chance of being exchanged with the corresponding segment in the other genome. Finally the new organisms are placed in the population, adjacent to one of their parents. Each time a sexual organism is placed in the environment they undergo migration the same way asexual organisms do. Sexual organism being placed in the environment there is a small random probability that it will migrate different deme and be placed there.

We measured three main metrics in all experiments: fitness, genetic diversity, and how many complex functions evolved. Fitness was a ratio of total cpu speed up and the time an organism took to make a copy of itself. We recorded fitness of the most abundant genotype at the end of each experiment, this genotype is referred to as the final dominant. Diversity was measured in terms of Shannon information entropy on the number of genotypes in the population. This equation is given below (equ 1), where p_i is the proportion of the ith genotype in the population. Populations that had more genotypes have higher information entropy, populations with fewer genotypes have lower information entropy. Finally, we counted how frequently XOR and EQU, the two most complex functions (Lenski et al 1999, Lenski et al 2003), were performed. Unstructured populations of digital organisms with this size and mutation rate will only rarely evolve all nine logical functions (Covert et al 2013, Covert and Wilke 2014).

$$H\left(genotypes\right) = -\sum p_i \log_2 p_i \qquad (1)$$

In all treatments resources were limited to an inflow of 10,000 units per update, distributed uniformly across the population so that each cell in the grid received one unit of resource. Each organism could use 1 unit of resource to perform one logical function during its life cycle. If no resource was

---

[1] Updates are a unit of time in Avida. Each update the population executes 30 CPU cycles per living organism. CPU cycles are awarded based on how many logical functions the organism completes successfully.
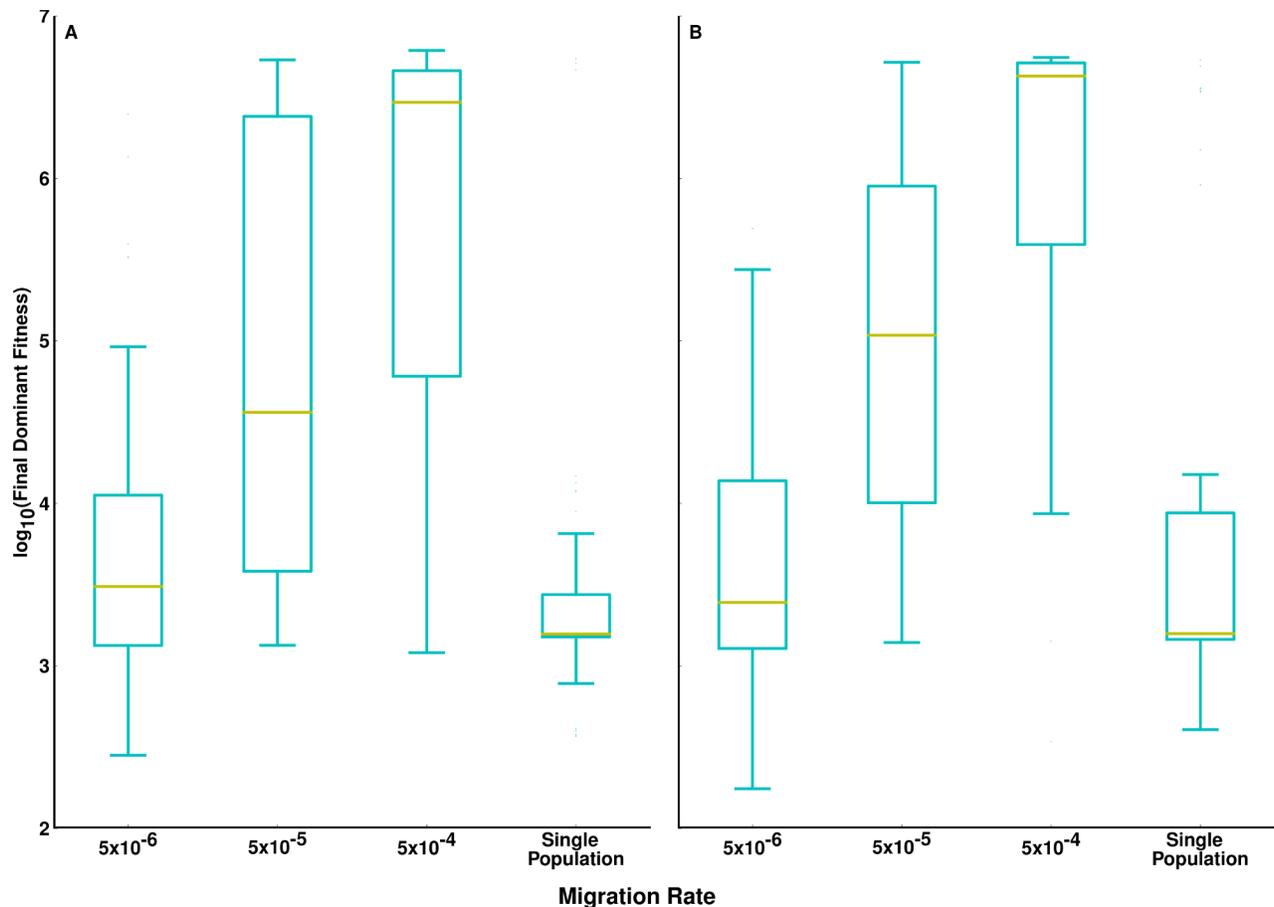
Figure 1 Log$_{10}$ Final dominant fitness at the end of each experiment for sexual (A) and asexual (B) experiments. All structured populations had higher median fitness than the single population controls. Differences in the $5x10^{-4}$ and $5x10^{-5}$ migration rates were significant, while the $5x10^{-6}$ migration rate was not significantly different from the control. See text for full statistical analysis.

available then no reward was granted for the function. Resources were spatially distributed such that each cell on the grid had its own pool of resource. Past works in Avida, have used globally distributed pools of resources (Chow et al 2006, Walker and Ofria 2012). Allocating resources to individual grid cells meant that organisms in different demes could not utilize the same pool of resources.

Results

Final dominant fitness from all treatments reveals a clear trend, in both sexual and asexual populations (Figure 1). At the highest migration rate, $5x10$^-4, both sexual and asexual populations evolve higher fitness than in any other treatment (Figure 1), we refer to this migration rate as the optimal migration rate. Both sexual and asexual treatments have significantly higher fitness at the optimal migration rate (p < 0.05 and p < 0.001 respective Mann-Whitney U-test for comparisons with all other treatments).

Fitness remains elevated in the other structured populations, however only the 10^-5 migration rates was significantly

higher than the control populations (both p<0.05 Mann-Whitney U-test). As migration rate decreased, so did fitness, suggesting that genetic drift began to take over in the individual demes as geneflow decreased. This results is constant with pervious results in structured populations (Covert and Wilke 2014), however the range of optimal migration rates is larger here than was perviously observed.

We also measured genetic diversity, as a function of information entropy on the number of genotypes (see methods). Genetic diversity was inversely correlated with fitness in all of the structured populations, both asexual and sexual (Figure 2). Sexual structured populations had significantly higher genetic diversity than asexual populations (all p<0.001, Mann-Whitney U-test). The asexual unstructured population had no significant difference in diversity with the optimal migration rate (p=0.155, Mann-Whitney U-test), but lower migration rates had significantly higher diversity (p<0.01, Mann-Whitney U-test). However, the sexual unstructured population had higher diversity than the sexual structured populations (p<0.05, Mann-Whitney U-test). Despite increased genetic diversity, sexual populations were
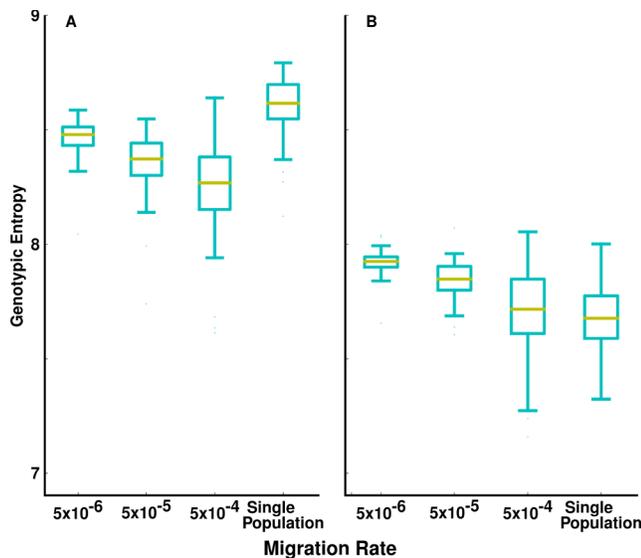
Figure 2 Genotypic entropy of the final population of sexual (A) and asexual (B) experiments. Structural populations had an inverse relationship with with migration rate. All sexual populations have significantly higher fitness than their asexual counter parts. The single population control had significantly higher diversity in the sexual case. The asexual control population had lower diversity or no significant difference in diversity from the asexual structured populations. See text for full statistical analysis.
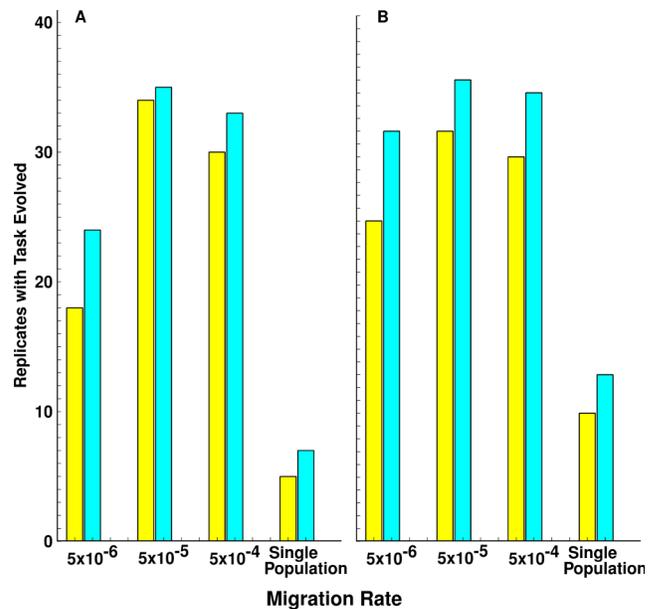

Figure 3 Complex function counts for all sexual (A) and asexual (B) treatments, showing the number of times of XOR (yellow) and EQU (cyan) evolved. All structured populations had significantly improved fitness over their controls. The lowest migration rate, 5×10^{-6}, had significantly fewer complex functions than the higher migration rates, but was still significantly higher than the single population control in both the sexual and asexual cases. See text for full statistical analysis.

unable to achieve higher fitness than their asexual counterparts, and the unstructured sexual population had lower fitness than all of structured treatments.

We measured how frequently the two most complex functions, XOR and EQU evolved in each treatment (Figure 3). Structured environments evolved complex function significantly more than the unstructured controls did (all p<0.001 Fisher's Exact test). Within just the structured treatments, asexual treatments at the two highest migration rates had no significant difference in the evolution of complex functions(XOR: p=0.79, EQU: p=1.0). The lowest asexual migration rate tested had fewer complex functions than the other two, but those differences were not significant (p=0.33 and p=0.55). The lowest asexual migration rate still evolved more complex function than the control (both p<0.05). Sexual structured populations followed a similar trend, with the two highest migration rates exhibiting high rates of task evolution and were not significantly different from one another (XOR: p=0.40, EQU: p=0.76). Even the lowest sexual migration rate produced more populations performing complex functions than the unstructured control.

While complex functions evolved at high rates over all migration rates, fitness remained relatively low at lower migration rates. This could indicate that high migration rates bring fit organisms to demes with pools of resources with greater frequency, keeping at least one subpopulation at high fitness at all times. Either way, structured populations had significantly improved evolution.

Discussion

Structured populations with limited resources demonstrate improved evolution over their unstructured counter parts. Both sexual and asexual structured populations exhibited improved fitness, genetic diversity and evolved more complex functions. Complex functions evolved frequently over a broad range of migration rates, even though resource limitations depressed overall fitness, particularly at lower migration rates.

Previous results in unlimited resource structured populations have shown an improved performance only at certain optimal migration rates (Covert and Wilke 2014). The optimal migration rate observed here was an order of magnitude higher than that of the previous study, indicating that the rate of resource inflow may determine which migration rate was optimal.

Other migration rates in unlimited environments were not significantly different from the control, or significantly lower than the control at very low migration rates. Structured populations evolved complex functions at a higher rate than the unstructured control, and at a broad range of migration rates. While overall fitness may be depressed due to a lack of resources, complex functions still evolved at a high rate. From this we may conclude that structured populations with limited resources are better able to adapt to their environment than

unstructured populations, or structured populations with unlimited resources.

Our populations with a simple structure and only a single limited resource evolved complex functions at extremely high rates. Evolving different complex functions in the same population is an extremely difficult problem, requiring organisms to maximize their evolutionary potential, while still exploiting already discovered fitness peaks. Evolutionary computation researchers have devised many ways of achieving this maximization through complex ecologies (Goings et al 2012) or through structured populations that impose artificial rules on migration between subpopulations (Hu et al 2005). Our results suggest that a balance between exploration and exploitation may be achieved by mirroring relatively straightforward structures and resource distributions inspired by natural systems.

These results could be altered by a number of factors. Lower resource inflow rates could result in reduced evolution. Here we ensured that every organisms could get rewarded for at least one function execution every update. At lower resource inflow rates selective pressure for complex functions would be reduced because rewards for evolving complex functions would become increasingly rare. At higher resource inflow rates populations would begin to evolve more like populations with infinite resources. Our work indicates that there is a range of optimal inflow rates, where a broad range of migration rates gives a distinct advantage to structured populations over unstructured populations.

Overall, structured populations are better able to adapt to limited resource environments than unstructured populations. This effect is detectable at a large range of migration rates and clearly indicates that resource limitation can dramatically improve evolutionary search in structured populations. This result addresses a long standing controversy in evolutionary biology, indicating that structured populations spread out over a large range may have improved adaptive ability. In addition, it provides further insight into methods that could be used to address difficult problems in computational evolution.

Acknowledgments

References

Belding T.C., 1995 The distributed genetic algorithm revisited. In Eschelman L. Ed., Proceedings of the Sixith International Conference on Genetic Algorithms,114-121, Morgan Kaufmann San Fransico.

Cantú-Paz, E. 1998 A survey of parallel genetic algorithms. Calculateurs Paralleles, Reseaux et Systems Repartis. 10, 141-171.

Chow S.S., Wilke C.O., Ofria C., Lenski R.E., and Adami C. (2004). Adaptive radiation from resource competition in digital organisms. Science 305:84-86.

Covert A.W., Carlson-Stevermer J., Derryberry D.Z., Wilke C.O. 2012. The Role of Deleterious Mutations in the Adaptation to a Novel Environment. In Artificial Life 13, eds Adami C, Bryson D, Ofria C, and Pennock RT (MIT Press) pp 27-31.

Covert A.W. and Wilke, C.O. (2014) Intermediate Migration Yields Optimal Adaptation in Structured, Asexual Populations. bioRxiv doi:10.1101/003897

Coyne, J. A., Barton, N. H., Turelli, M., 1997. Perspective: A Critique of Wright's Shifting Balance Theory of Evolution. Evolution. 51: 643-671

Coyne, J. A., Barton, N. H., Turelli, M., (2000) Is Wright's Shifting Balance Process Important in Evolution? 54:306-317

Fernández, F., Tomassini, M., & Vanneschi, L. 2003 An empirical study of multipopulation genetic programing. Genetic Programing and Evolvable Machines 4, 21-51.

Gerrish, P. J. and Lenski, R. E., 1998. The Fate of Competing Beneficial Mutations in an Asexual Population. Genetica, 102/103: 127-144.

Goodnight, C. J., and M. J. Wade. 2000. The ongoing-synthesis: a reply to Coyne et al. (1999). Evolution 54:317-324.

Kryazhimskiy, S., Rice, D. P., Desai, M. M., 2012. Population Subdivision and Adaptation in Asexual Populations of Saccharomyces Cerevisiae. Evolution, 66:1931-1941.

Lenski R.E., Ofria C., Collier T.C., Adami C. 1999. Genome Complexity, Robustness and Genetic Interactions in Digital Organisms. Nature 400:661-664.

Lenski R.E., Ofria C., Pennock R.T., Adami C. 2003. The Evolutionary Origin of Complex Features. Nature. 423:139-144.

Lin, S. C., Punch, W. F., & Goodman, E. D. 1994 Coarse-grain parallel gentic algrorithms: categorization and new approach. Proceedings of the Sixth IEEE Symposium on Parallel and Distributed Processing. 28-37, IEEE

Miller, C. R., Joyce, P., Wichman, H.A., 2011. Mutational Effects and Population Dynamics During Viral Adaptation Challenge Current Models. Genetics, 187:185-202.

Misevic D., Ofria C. A., Lenski, R. E., 2006 Sexual reproduction reshapes the genetic architecture of digital organisms. Proc. R. Soc. London B 273, 457-464.

Moore, F. B. G. and Tonsor S. J., 1994. A Simulation of Wright's Shifting Balance Process and the Three Phases. Evolution, 44:69-80.

Ofria, C. and Wilke, C.O., 2004. Avida: A Software Platform for Research in Computational Evolutionary Biology. J. Artif. Life, 10:191-229.

Phillips, P. C. (1996) Waiting for a Compensatory Mutation: Phase Zero of the Shifting-Balance Process. Genet. Res., 67:271-283

Wade, M. J., Goodnight, C. J., Perspectives: The Theories of Fisher and Wright in the Context of Metapopulations: When Nature Does Many Small Experiments. Evolution, 52: 1537-1553

Walker B and Ofria C (2012) Evolutionary Potential is Maximized at Intermediate Diversity Levels, Proceedings of the 13th International Conference for Artificial Life, East Lansing, MI.

Wright, S., 1932. The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. Sixth International Congress on Genetics, Ithica, New York, 356-366.