ウェブサイト公開用

（様式 10）
（海外特別研究員事業）

令和　2 年　4 月　29 日

海 外 特 別 研 究 員 最 終 報 告 書

独立行政法人日本学術振興会 理事長　殿

採 用 年 度　2018
受 付 番 号　60528

氏　　名　井筒 弥那子

（氏名は必ず自署すること）

海外特別研究員としての派遣期間を終了しましたので、下記のとおり報告いたします。
なお、下記及び別紙記載の内容については相違ありません。

記

1. 用務地（派遣先国名）用務地：　イーストランシング　（国名：　アメリカ合衆国　）

2. 研究課題名（和文）※研究課題名は申請時のものと違わないように記載すること。
適応進化における集団内多型と新規変異の役割の実験科学的評価

3. 派遣期間：平成　30　年　4　月　1　日　～　令和　2　年　3　月　31　日

4. 受入機関名及び部局名
ミシガン州立大学　微生物分子遺伝学部

Some basic evolutionary issues can lead to misunderstandings and confusion, even among experts. One such issue concerns the **contributions of standing genetic variation versus new mutations to the process of adaptation by natural selection**. It's a vexing problem because standing variation and new mutations are closely related and there's no single "right" answer in terms of relative contributions. But the ways we do science—both conceptually and empirically—often lead us to emphasize one or the other source of genetic variation. In the *E. coli* long-term evolution experiment (LTEE) conducted by Dr. Richard Lenski, Michigan State University, USA, for example, new mutations have been emphasized because the replicated populations were founded from a common ancestral strain to ensure independence and examine the repeatability of evolution [Lenski et al., 1991; Woods et al., 2006; Tenaillon et al., 2016; Lenski, 2017], so all of the genetic variation was generated by new mutations. Hence, there was no standing variation at the start, although it accumulated as mutations arose and spread in the population.

Much of the work in the field of experimental evolution follows the same mutation-driven strategy including most studies of bacteria and yeast as well as digital organisms [Rainey & Travisano, 1998; Lang et al., 2013; Lenski et al., 2003]. **However, that is not the approach used in many other systems**, including fruit flies [Graves et al., 2017; Burke et al., 2010], for two largely practical reasons. First, quantitative genetics theory [Falconer & Mackay, 1996; Barton & Keightley, 2002], which was developed for sexual plants and animals, is based on standing genetic variation, whereas the input from new mutations is typically ignored or abstracted. Second, the long generation time and small population size in these systems make evolution experiments with isogenic inbred lines impractical, and so researchers usually start with outbred lines [Rose, 1984; Scarcelli & Kover, 2009; Long et al., 2014]. These differences in perspectives and approaches have sometimes been a source of confusion.

In this project, **I directly compare the rates of adaptation based on standing genetic variation versus new mutations**, using the LTEE lines as ancestors and challenging them with a **new environment**. At one extreme, I prepared populations with high initial variation by mixing multiple independently evolved LTEE populations from 50,000 generations; at the other extreme, each population was founded from an individual clone from the same LTEE populations, such that there is no initial within-population variation. All of the populations then evolved for 2,000 generations in the new environment.

First, I performed some preliminary experiments to decide experimental conditions. I checked the marker status of the LTEE populations and clones from generation 50,000. At the beginning of the LTEE, Dr. Lenski prepared 12 populations among which half of them derived from one Ara+ strain and the other half derived from one Ara- strain [Lenski et al., 1991]. Those ancestral Ara+ and Ara- strains are identical except for one selectively neutral marker mutation in *araA* gene and those two strains can be distinguished on tetrazolium arabinose (TA) plates. This experimental design has enabled us to check possible contaminations during the long-running experiments and estimate fitness of the evolved lines by competition assays. However, some of the LTEE strains have evolved phenotypes that show unclear colonies on TA plates and there is also one LTEE strain which has evolved the ability of uptake of citrate, which is present in the media as a chelating agent, as a carbon source [Blount et al., 2008; Blount et al., 2012]. I need to use strains showing clear distinguishable colonies on TA plates especially for measuring fitness and strains exploiting one carbon source in the media to make sure the evolving lineages compete against each other for the only carbon source. Hence, I checked the marker status of 12 LTEE populations, isolated clones and excluded lines that didn't show the necessary phenotypes for this project. I also checked the growth rate of LTEE lines in the media with different carbon sources. The carbon source in this new experiment should be different enough from glucose, which is the only carbon source in the original LTEE environment, because I intend to assess standing genetic variation and new mutations as a fuel for adaptive evolution and it is highly possible that the ancestral populations carry adaptive alleles from the beginning if I use carbon source which is similar to glucose. A former postdoc, Dr. Christopher J. Marx, examined the growth rate of LTEE clones with various carbon sources [Leiby & Marx, 2014]. Based on their results, I chose several

candidates for the carbon source in the new environment and measure the growth rate and colony forming unit (CFU) of all of the LTEE populations and clones from generation 50,000 that show clear colonies on TA plates. I chose generation 50,000 because we have already sequenced whole population genomes of all of the 12 LTEE populations and genomes of 2 isolated clones each from the 12 populations. Among the 12 populations, Ara+1, +3, +4, +6, and -2 didn't show clear colonies on TA plates and Ara-3 has evolved the ability to exploit citrate, so I excluded those 6 strains from the subsequent growth measurement.

I measured growth rate of 6 LTEE lines, Ara-1, +2, -4, -5, +5, and -6, on mannose, lactose and D-serine media, confirmed those 6 LTEE lines grew well on all of the three carbon sources and finally chose D-serine as an only carbon source in this new evolution experiment. D-serine is an amino acid found in soils and aquatic environments, although most serine in those environments is the isomeric L-serine used in proteins [Kubota et al., 2016]. D-serine is also abundant in mammalian brains [Hashimoto et al., 1992] and urine [Huang eta., 1998; Anfora & Welch, 2006], and it has a bacteriostatic role in the urinary tract [Sasabe & Suzuki, 2018]. Although D-serine has attracted considerable medical interest, little is known about how it inhibits bacterial growth while also serving as a carbon source [Connolly etal., 2016]. There is substantial variation in growth rate among the LTEE lines in D-serine media [Leiby & Marx, 2014], making it an interesting new environment. I set the concentration such that the bacteria reach the same stationary-phase density as in the LTEE, while the use of D-serine as the source of carbon and energy makes the environment substantially different from that of the LTEE.

This experiment has 4 treatments (Fig. 1): (A) each population starts from a single clone sampled from one LTEE population; (B) each population starts from a heterogeneous LTEE population; (C) each population starts as an admixture of treatment A populations; and (D) each population starts as an admixture of treatment B populations. There are 72 populations in total: treatments A and B have 6 source populations, each with 3-fold replication; treatments C and D use the same source populations, each with 18-fold replication. **Populations in treatment A start from a single genotype and thus depend entirely on new mutations for adaptation, while those in D start with the maximum diversity; B and C provide intermediate levels of initial genetic variation.** The culture conditions are almost the same as those used in the LTEE [Lenski et al., 1991]. Briefly, I transfer 100 ul of the culture into 9.9 ml of minimal D-serine media in glass text tubes daily, mix them well, and incubate them at 37 ℃ without shaking for 22-26 hours. The 100 ul of culture contains about 5×10^6 cells and they grow to 5×10^8 cells at the end of the cycle. This 100-fold dilution gives about 6.64 generations per day and we have found that 100-fold daily dilution protocol maximizes the rate of adaptation in our system (Izutsu et al., in preparation) though there are some differences in experimental condition between this project and the previous work.

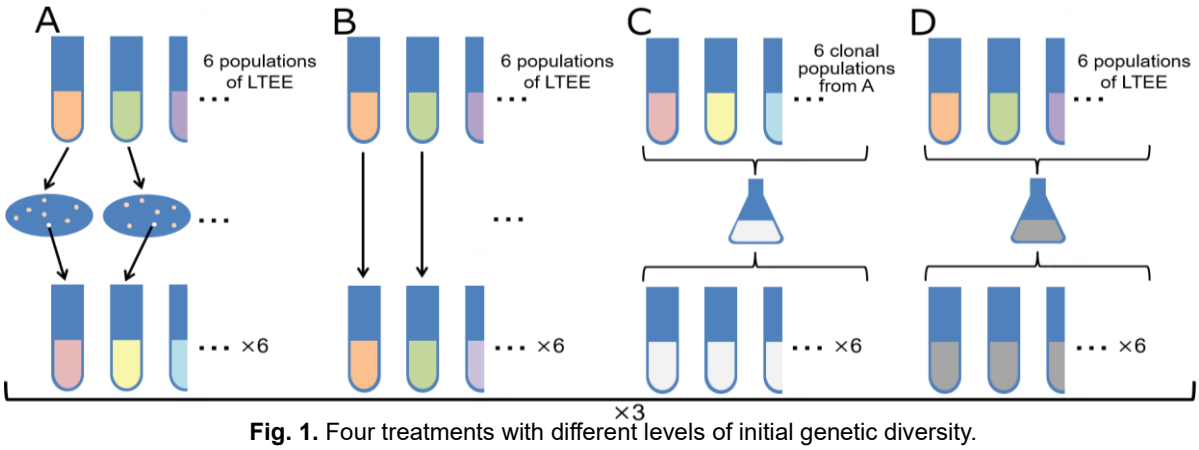

Fig. 1. Four treatments with different levels of initial genetic diversity.

During the evolution experiment, the populations had been propagated in the new environment with D-serine for 2,000 generations, which took 300 days given 1:100 serial dilutions each day. Samples were frozen at every 100 generations and used for subsequent analyses such as fitness measurement or genome sequencing. The evolution experiment reached 2000 generations last fall. I have plated the populations of treatment C & D onto TA plates on a regular basis. By tracing the Ara- and Ara+ marker ratio, I could observe the selective sweep and that helped us estimate the effect of standing variation. I expected that the populations carrying

standing variation would experience rapid selective sweep right after the selection started because the pre-existing beneficial variation would expel the other variation within the population. Replicate populations in **treatment C showed significant convergence in terms of Ara marker divergence** (upper panel in Fig. 2). Intriguingly, many populations in treatment D had not experienced any hard selective sweep at until day 200 and several populations even kept both marker lineages at day 300 though we also observed significant convergence in marker ratio at around day 10 (lower panel in Fig. 2). This indicates **populations in treatment D have experienced long competition among lineages with different beneficial mutations**. The previous works in Dr. Lenski's lab revealed that the frequency-dependent interactions between lineages could cause the long coexistence of different adaptive lineages in a population [Maddamsetti et al., 2015; Ribeck & Lenski, 2015]. I also have plated all of the 72 populations onto TA plates every 100 generation to check possible contamination. I observed emergence of unusual colonies in many lines. Most of them are extremely tiny colonies which looked white but turned to red when I streaked them onto TA plates separately. Thus, the marker divergence might not strictly reflect the actual amount of each marker lineages. However, I did not observe any critical contaminations during this 300-day-long evolution experiment.

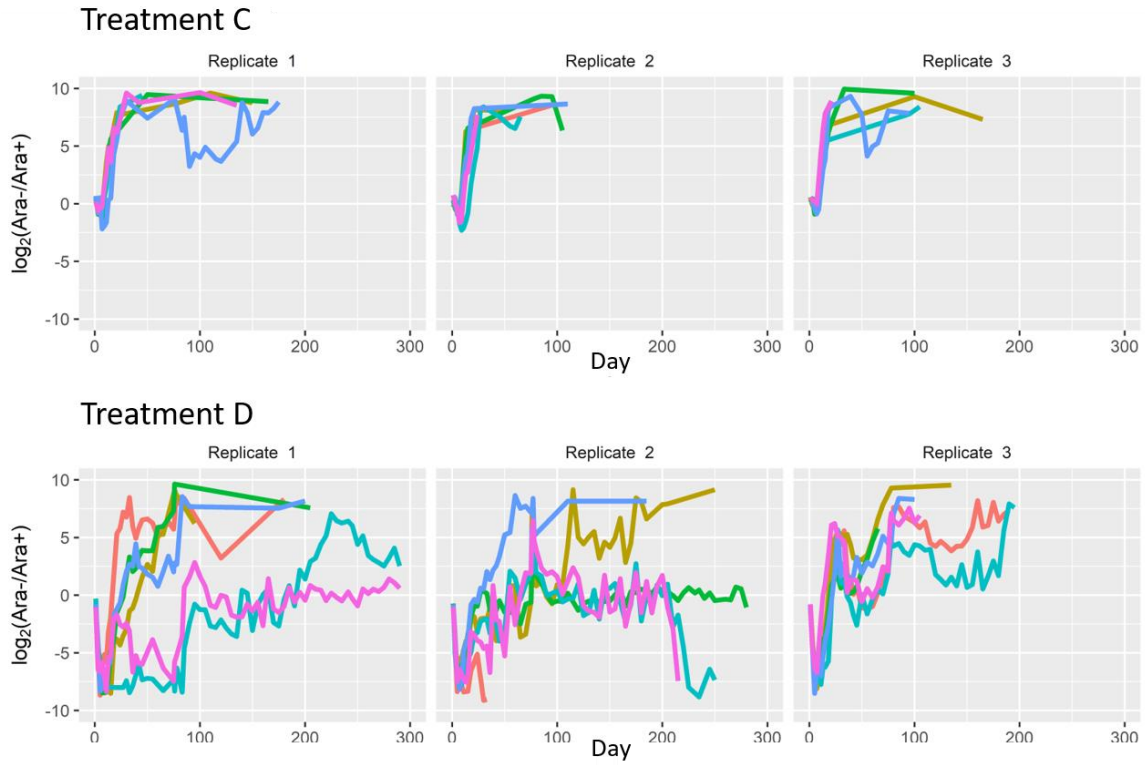

Fig. 2. Marker divergence in treatment C & D during 300 days of evolution experiment

After checking contamination, I measured fitness of the evolved lines. As I mentioned above, populations in treatment D still have sustained both marker lineages at most of the time point so we cannot measure fitness of populations. Instead, I isolated clones from each evolved population, checked the marker status, and competed them with the common ancestor strains. We measure fitness by mixing an evolved line and a common ancestor strain that show the opposite Ara marker usually with 1:1 ratio, let them grow in the media for 1 to several days, and calculate fitness based on the number of colonies on TA plates at day 0 of the competition and that at later day. When I tried to measure fitness of evolved clones with 1:1 starting ratio, it turned out the colonies from the common ancestor competitior were few on TA plates even day 1 of the competition assays, indicating the fitness of these evolved lines is extremely high. If the number of colonies of the competitor is too small, we cannot assure the accuracy of the fitness measurement. There are several solutions for measuring such high fitness. For example, changing the common competitors to the strains that have relatively higher fitness than the ancestors that we used, plating the culture before 24 hours, or changing the starting ratio. The common ancestors we used seemed to show high enough fitness in D-serine media according to the previous works [Leiby & Marx, 2014], thus it is possible that we will face the same issue even we change the competitior strains. There are also several advantage of using the common ancestors that we have used in LTEE. Plating the culture before 24 hours makes experiments complicated and 24 hours in the

D-serine media is the environment they have been exposed and adapted to, so checking the fitness before 24 hours is not an ideal solution. Therefore, I chose to change the starting ratio. Based on the preliminary results, I decide to mix them with 1:4 ratio at the beginning of the competition assays.

Fig. 3 shows the results of preliminary fitness assays and **I confirmed significantly large fitness increases in D-serine media**, greater than those seen in 50,000 generations of the LTEE which is fitness < 2 [Wiser et al., 2013]. I could not measure fitness of some lines that showed too small colonies to count. Nonetheless, the preliminary results suggest **higher mean fitness (gray bars) in treatment A than in B**, which contradicts the general prediction: fitness gain would be higher in the population carrying standing variation than that without initial variation. When we proposed this project to an internal grant in our institute, one reviewer criticized it, saying that the expected results were almost guaranteed. **This result also implies a rugged fitness landscape with widely divergent slopes or even multiple local peaks.**

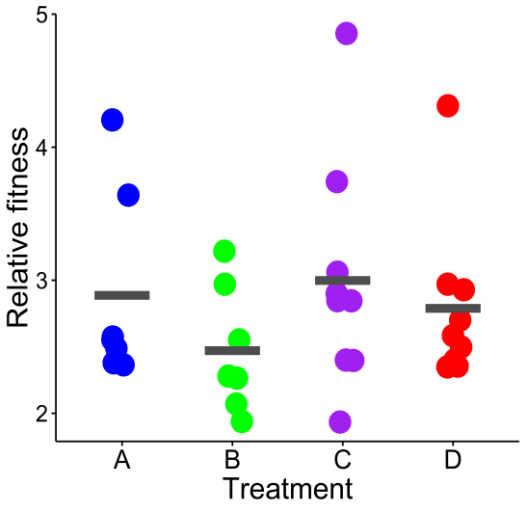

Fig. 3. Fitness of clones at generation 2,000.

I was isolating clones from different time points to examine time-series change in fitness. Unfortunately, we needed to stop experiments because of the current COVID-19 outbreak. I am going to finish isolating clones from generation 500, 1000, and 1500 as soon as we open up our institute and am going to obtain more fitness data. I am also going to perform computer simulations as a collaboration with a graduate student in Lenski's lab to better interpret our data.

I talked about this project at seminars in Michigan State University several times and obtained many fruitful advice from colleagues in various field of science. I am going to talk about this project in Evolution meeting 2020 or an internal congress of the BEACON center for study of evolution virtually held on August 2020.