Biocatalysis, Protein Engineering, and Nanobiotechnology **Biotechnology and Bioengineering**
DOI 10.1002/bit.25111

Simultaneous improvement of specificity and affinity of aptamers against *Streptococcus mutans* by *in silico* maturation for biosensor development[†]

Nasa Savory[1,3], Yayoi Takahashi[2,3], Kaori Tsukakoshi[1], Hijiri Hasegawa[1], Madoka Takase[1], Koichi Abe[1], Wataru Yoshida[1], Stefano Ferri[1], Shizuko Kumazawa[2], Koji Sode[1] and Kazunori Ikebukuro[1]*

[1] Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-26 Naka-cho, Koganei, Tokyo 184-8588, Japan

[2] Devices Development Center, TDK Corporation, 2-15-7 Higashi-Ohwada, Ichikawa, Chiba 272-8558, Japan

[3] These authors equally contributed to this work.

* Corresponding author:

Address: Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo, 184-8588, Japan

E-mail: ikebu@cc.tuat.ac.jp

Tel/Fax: +81 42 388 7030

Short Running Title: Evolution of aptamer specificity and affinity

[†]This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/bit.25111]

Additional Supporting Information may be found in the online version of this article.

Received May 26, 2013; Revision Received August 27, 2013; Accepted September 3, 2013

Abstract

In silico evolution with an *in vitro* system can facilitate the development of functional aptamers with high specificity and affinity. Although a general technique known as systematic evolution of ligand by exponential enrichment (SELEX) is an efficient method for aptamer selection, it sometimes fails to identify aptamers with sufficient binding properties. We have previously developed *in silico* maturation (ISM) to improve functions of aptamers based on genetic algorithms. ISM represents an intelligent exploitation of a random search within a defined sequence space to optimize aptamer sequences and improve their function of interest. Here we demonstrated a successful application of ISM of aptamers to simultaneously improve specificity and affinity for *Streptococcus mutans* with discovery of a core sequence, which was required to form a polymerized guanine quadruplex structure for target binding. We applied ISM to aptamers selected by whole-cell SELEX and identified an aptamer with up to 16-fold improvement in affinity compared to its parent aptamers, and specificity was increased to show 12-fold more binding to *S. mutans* than to *Lactobacillus acidophilus*. Furthermore, we demonstrated a specific flow-through detection of *S. mutans* at a concentration range of 1×10^5–10^8 CFU/mL using the evolved aptamer immobilized on gold colloids.

Keywords

Aptamers, *Streptococcus mutans*, Cell-SELEX, *in silico* maturation, flow-through test

Introduction

Aptamers are single-stranded oligonucleotides that form unique three-dimensional structures and interact with target molecules with high affinity and specificity. The selective interactions between aptamers and diverse targets have been exploited for different applications. With advantages in targeted therapy, imaging, and diagnosis, aptamers are being readily considered as potential targeting ligands (Fang and Tan 2010). Aptamer-based biosensors for specific detection of pathogenic microorganisms are among the potential applications (Torres-Chavolla and Alocilja 2009).

Typically, aptamers are identified from a random oligonucleotide library through an *in vitro* selection known as systematic evolution of ligands by exponential enrichment (SELEX) (Ellington and Szostak 1990; Tuerk and Gold 1990). In addition, modified SELEX techniques targeting live cells (Cell-SELEX) including mammalian (Shangguan et al. 2006) and bacterial cells (Chen et al. 2007) have been used to select aptamers against membrane-associated proteins in their native conformations even without prior knowledge of target proteins. Despite these advantages of SELEX, there are several drawbacks. For example, iterative PCR may result in sequence bias of library oligonucleotides and fail to recover aptamers (Schutze et al. 2011). Although aptamer selections relying on high-throughput sequence analysis have been performed, it is known even a reduced number of PCR process may critically affect aptamer identification (Cho et al. 2010). Furthermore, random oligonucleotide pools are known to be not structurally diverse and favor simple topological structures (Gevertz et al. 2005). Considering these problems, Mayer *et al*. provided a conservative estimate of approximately 50% for the success rate of SELEX (Mayer et al. 2010). Thus, an additional approach is required for development of aptamers with high affinity and specificity.

Computational approaches for screening and optimizing aptamer sequences would counter the loss of sequence diversity and facilitate development of functional aptamers. *In silico* evolution with an *in*

vitro system based on genetic algorithms (GAs) is a promising strategy for exploring the sequence-fitness landscape (Knight et al. 2009; Platt et al. 2009) and improve specific functions of aptamers (Ikebukuro et al. 2005; Noma and Ikebukuro 2006; Noma et al. 2006; Nonaka et al. 2013; Savory et al. 2010; Savory et al. 2013). GA represents an intelligent exploitation of a random search within a defined search space to solve complex and multidimensional problems (Goldberg 1989; Holland 1975). A directed evolution-based strategy is known as an effective approach for optimizing aptamer sequences to a specific scaffold of aptamers instead of rational design (Ahmad et al. 2012). In aptamer selection, we have established a method based on GAs, designated as *in silico* maturation (ISM) to identify aptamers with improved features such as affinity (Nonaka et al. 2012; Savory et al. 2010), specificity (Savory et al. 2013) and enzyme-inhibitory activities (Ikebukuro et al. 2005; Noma and Ikebukuro 2006; Noma et al. 2006).

ISM is a directed evolutionary technique employing an iterative process of *in silico* sequence modifications followed by *in vitro* evaluation. In this method, all candidate sequences are produced *in silico* and they are chemically synthesized for *in vitro* evaluation without PCR amplification, thus PCR bias does not issue in the ISM process. Therefore, this method progressively expands the sequence space in which aptamers are screened and facilitate optimization of aptamer sequences. Moreover, it can improve aptamer functions other than affinity if the functions of interest can be evaluated in *in vitro* or even *in vivo* experiments. In contrast, the SELEX technique selects aptamers based only on affinity.

In this study, we demonstrated a successful application of ISM to aptamers for simultaneously improving specificity and affinity for *Streptococcus mutans*, a cariogenic pathogen and the predominant etiological agent of dental caries and the cause of increased global healthcare costs (Hamada and Slade 1980). We applied ISM to aptamers selected by Cell-SELEX and isolated an aptamer with improvement of affinity and specificity over *Lactobacillus acidophilus*. This is the

primary application of ISM for simultaneously improving two functions of aptamers. In addition, we proposed the first example of ISM to discover a schema, which is known as a milestone of GA process. We identified a schema as a certain sequence required for aptamer binding, through iterative process of GAs. The discovery of schema and subsequent optimization of aptamer sequences facilitated the evolution of high fitness sequences based on GAs. Furthermore, we constructed a flow-through detection platform using the improved aptamer for specific detection of *S. mutans* by aptamer-immobilized gold colloids.

Materials and Methods

Bacterial strains and cultures

S. mutans ATCC 25175 and JSM 5175 were purchased from Kanto and Riken BRC, respectively. Isolated strains of *S. mutans* (F1, F2, F3, and F4) were obtained from the culture collection of TDK Corporation. *Streptococcus sobrinus* ATCC 334478 and JSM5176 were purchased from Summit Pharmaceuticals International and Riken BRC, respectively. *Streptococcus salivarius* JCM5707, *Streptococcus sanguinis* JCM 5708, and *Streptococcus anginosus* JCM12993 were purchased from Riken BRC. *Streptococcus mitis* ATCC 49456 and *Streptococcus oralis* ATCC 35037 were purchased from Summit Pharmaceuticals International. *Lactobacillus acidophilus* ATCC 4356 was purchased from Kanto. All cells were grown in BBL brain heart infusion broth (Becton Dickinson) for 20 h at 37 °C, and 10% glycerol stocks were prepared and stored at −80 °C until use. The cells were washed in TBS buffer (10 mM Tris-Cl, 150 mM NaCl, 5 mM KCl, pH 7.5) before all experiments.

DNA library

DNA oligonucleotides were synthesized by Life Technologies. Single-stranded DNA (ssDNA) libraries consisting of a 30-nt randomized region (N30) or 60-nt randomized region (N60) flanked by 18-nt

primer sequences were used for Cell-SELEX (Table S1). To fold aptamer structures, the libraries and identified aptamers were subjected to heat treatment at 95 °C for 10 min and ramped to 25 °C over 30 min before all experiments.

Aptamer selection

One microliter each of *S. mutans* ATCC 25175 (1×10^{10} CFU/mL) and *L. acidophilus* (1×10^{8} CFU/mL) prepared in TBS buffer were spotted on a single piece of Hybond-ECL nitrocellulose membrane (GE Healthcare). The library [final concentration (f.c.) 1 μM for the first round and 40 nM for subsequent rounds] was incubated with the membrane for 1 h at room temperature (r.t.). After washing, the spot area of *S. mutans* was excised and library oligonucleotides were extracted by phenol–chloroform. The extracted oligonucleotides were amplified by PCR using 5′-fluorescein-labeled forward primer, 5′-biotin-labeled reverse primer (Table S1), and AmpliTaq Gold DNA polymerase (Life Technologies). To prepare the ssDNA library for the following round, 5′-biotin-labeled reverse strands were removed using avidin agarose beads (Thermo Scientific). From the third to fifth rounds of Cell-SELEX using the N60-L1 library (Table S1), aptamer selection was performed for *S. mutans* without immobilizing *L. acidophilus* on the nitrocellulose membrane. After four and five rounds of aptamer selection, library oligonucleotides were cloned and sequenced.

Aptamer blotting assay

Bacteria cells (f.c. 1×10^{9} CFU/mL) were incubated with 5′-fluorescein-modified aptamer (f.c. 1 μM) and 1/500 diluted anti-FITC antibody conjugated with horseradish peroxidase (HRP) (Dako, Denmark) in TBS-T [TBS plus 0.05% (v/v) Tween 20] buffer for 30 min at r.t. Following washing, cells were spotted on a nitrocellulose membrane and HRP chemiluminescence was detected by Typhoon 8600 (GE Healthcare) using ECL Plus Western Blotting Detection Reagents. HRP chemiluminescence signal intensity was assessed using ImageQuant TL software (GE Healthcare).

In silico maturation of affinity and specificity

For the first cycle of ISM, ten partial sequences of Cell-SELEX-selected aptamers were chosen as parents (P) (Table S2) for *in silico* sequence modification. The parent sequences were crossed over with a different parent sequence at a random point following random point mutations with a mutation rate of 10%. These sequence modifications were applied to produce a set of 16 sequences as the first generation (G1). The G1 sequences were synthesized for *in vitro* evaluation, and their binding abilities to *S. mutans* ATCC 25175 and *L. acidophilus* cells were assessed by aptamer blotting assays as described above. As a control, one of parent aptamers SMa#3-6-P1 (5′-ATACCAGCTTATTCAATTGGGGGGAGGGGGGGTTGGGCTGTGGGGCGGTTAGCCGC-3′) was assessed at the same time, and binding signal of each sequence was normalized by that of SMa#3-6-P1 as 1.00. This aptamer was chosen due to its relatively high affinity and specificity for *S. mutans* among parent sequences. To produce the G2 sequences, the top eight G1 sequences showing highest binding ability to *S. mutans* were chosen and sequence modification was performed in the same manner. Then 20 G3 sequences were produced from the top 10 G2 sequences.

To improve binding specificity of aptamers for *S. mutans*, we set both binding ability and selectivity to *S. mutans* ATCC 25175 over *L. acidophilus* as fitness in ISM to produce the G4 and later generations. Binding selectivity was evaluated by the aptamer blotting assays as a ratio of signal intensity for *S. mutans* to that for *L. acidophilus*. To increase sequence diversity, the G4 sequences were produced from the top three sequences of each of P, G1, and G3 that showed highest selectivity to *S. mutans*. P and G1 as well as G3 sequences were used to generate G4 to expand sequence diversity. Then G5 sequences were produced from the top two of G3, the top three of G4, and two SELEX-selected sequences. The SELEX-selected sequences were used to further broaden the sequence diversity. The G5, G4, G3, parents, and SELEX-selected sequences were used to generate 32 sequences as G6. To

produce sequences from G6 to G9, the mutation rate was decreased to 5% from the 10% used for G1 to G5.

After the sixth generation of ISM, the *in silico* sequence modification process was altered to a shuffling approach. From G7 to G9, each sequence was divided into three blocks based on homology search to find conserved sequence blocks, and these blocks were shuffled between three randomly paired sequences (Fig. 1). To characterize a larger population of sequences, G9 was produced as a set of 98 sequences. To produce G10 to G11 sequences, the core sequence (5′-GGGGGGAGGGGGGG-3′), which was conserved among G9 sequences, was fixed as one block and the other two blocks were shuffled between three different sequences. Point mutations were then introduced into each block with a mutation rate of 15%. Finally, two aptamers showing the highest specificity and affinity were identified in the G11. All sequences produced during ISM are shown in Table S3.

Microscopy bead assay

Avidin agarose beads were incubated with 5′-biotin-modified aptamer SMa#G6-25 (5′-GGATCTGGTTAGTGAAGAGGGGGGAGGGGGGGTTGGGCTGTGGGGCGGTTAGTCGT-3′) (f.c. 14 μM) in immobilization buffer (40 mM Tris-Cl, 4 mM EDTA, 700 mM NaCl, pH 8.0) for 90 min at r.t. After washing in binding buffer [50 mM citrate, 5 mM KCl, 0.05% (v/v) Tween-20, pH 3.5], aptamer-immobilized beads were incubated with 5×10^7 CFU/mL *S. mutans* ATCC 25175 or *L. acidophilus* in binding buffer for 30 min at r.t. After wash in binding buffer, beads were suspended in binding buffer and examined by microscopy.

Flow-through test using aptamer-immobilized gold colloids

5′-Thiol-modified aptamer SMa#G11-5 (5′-ATACATCTTAAGTCTCGTGGGGGGAGGGGGGGTTGGTGGGCTTT-3′) was immobilized on

colloidal gold with 40 nm diameter (Funakoshi) by incubating overnight at r.t. in the presence of 5 mM DTT. Gold colloids were incubated with 10% (w/v) BSA for 3 min at r.t. The SMa#G11-5 gold colloid conjugates were washed in 5 mM phosphate buffer (pH 7.0). The conjugates were then added to *S. mutans* ATCC 25175 (1.0×10^8 CFU/mL to 5.0×10^4 CFU/mL) in PBS-T [5 mM phosphate buffer plus 0.05% (v/v) Tween 20] buffer. After mixing, the solution was immediately blotted on UltraBind affinity membrane (Pore size; 0.45 μm) (Pall) using a dot blotter (510 μL/spot). Spot intensity was assessed with ImageQuant TL software.

Results and discussion

Cell-SELEX for selection of aptamers

To identify DNA aptamers against *S. mutans* cells, we performed five individual Cell-SELEX experiments. Libraries with N30 or N60 sequences were used to vary the potential diversity. During selection, *L. acidophilus,* the most common bacterial species in the human mouth, was used as a competitor to eliminate aptamers that cross-react with non-target bacteria. After four and five cycles of Cell-SELEX, we identified aptamers that bound to *S. mutans* cells (Table S4), although they exhibited broad specificity for other bacteria including *L. acidophilus* in aptamer blotting assays (Ogasawara et al. 2007) (data not shown). This suggests that Cell-SELEX-selected aptamers bind to surface-associated molecules or common structures present on bacterial cells, even though we performed competitive screening with the aim of identifying aptamers specific to *S. mutans*. It should be noticed that the selected aptamers showed significant binding to *L. acidophilus* as well as *S. mutans*. It is known Cell-SELEX had difficulty in finding specific aptamers against expected target cells in a particular case (Kunii et al. 2011). For this reason, another approach to effective development of aptamers was required.

In silico maturation for improvement of affinity and specificity

We performed ISM to improve both affinity and specificity for *S. mutans*. The ISM was performed with three *in silico* sequence modification procedure (Fig. 1) to explore a wide sequence space for discovering a core sequence required for aptamer binding, following sequence optimization for the core sequence. To produce the first to sixth generation sequences (G1 to G6), we performed ISM to explore a wide sequence space based on sequence crossover and point mutations. From G7 to G9, we carried out shuffling of conserved sequence blocks other than crossover to optimize aptamer sequence and to discover a core sequence that is involved in binding ability of aptamers. Then, we further optimized aptamer sequences from G10 to G11 by shuffling sequence blocks with fixing the core sequence that was discovered by G9.

To carry out ISM of aptamers, ten partial sequences of Cell-SELEX-selected aptamers were chosen as parents (Table S2). Those partial sequences were designed based on the secondary structure prediction by M-fold (Zuker 2003) and sequence alignment. We designed partial sequences as those folding into the predicted secondary structures and also as sequences conserved among identified aptamers. Then, those partial sequences showing higher binding signals to *S. mutans* compared with full-length aptamers were chosen as parents to initiate ISM. We expected ISM started from parents with different lengths would optimize sequence length as well as sequence itself to show high fitness.

We generated a set of 16 sequences as the G1 from the parents by *in silico* sequence modification with crossovers and point mutations (Fig. 1). Based on *in vitro* assays, sequences showing the highest binding ability were chosen to produce a subsequent set of aptamers for further generations until G3. It should be notice that the G1 aptamers showed decreased binding signals compared to parent aptamers. We presumed that this was because *in silico* sequence modification process produced diverse sequences as we expected to search the sequence-fitness space avoiding local optima (Fig. 2A). It is

widely known that a GA inclines to locate populations to local optima at its early process, and it would be usually difficult to achieve further improvement. Therefore, we expected further process of our ISM approach would be able to find better aptamer candidates based on GAs avoiding local optima.

For producing G4 sequences, we selected aptamer sequences showing higher binding specificity for *S. mutans* than for *L. acidophilus* to improve specificity of aptamers to *S. mutans* because the G3 aptamers displayed binding to *L. acidophilus* as well as *S. mutans*. This selection process was expected to function for directed evolution of aptamers with high affinity and specificity for *S. mutans*. In addition, we used not only the immediate previous generation, but also several earlier generations for generating offspring. We expected these parent selections to increase sequence diversity for exploring a wide sequence space. Binding specificity for *S. mutans* was evaluated as a ratio of signal intensity for *S. mutans* to that for *L. acidophilus* by aptamer blotting assays. By the G6 of ISM, aptamers showing higher affinity and specificity for *S. mutans* were converged into sequences with a guanine-rich sequence block (Table S3) suggesting guanine-rich sequences were required for binding to *S. mutans*.

Therefore, we altered the process of sequence modifications to a shuffling approach based on evolution-mimicking algorithms reported previously (Ikebukuro et al. 2005; Noma and Ikebukuro 2006; Noma et al. 2006) to further optimize aptamer sequences and discover a specific sequence motif required for aptamer binding. Shuffling was performed with homology search to find sequence blocks conserved among different aptamers. Each block was then shuffled among three different sequences following point mutations (Fig. 1).

In this way we obtained G9 aptamers that contained guanine rich sequence portions flanked by variable sequences. Therefore, we considered those guanine rich portions were required for aptamer binding and we investigated binding of some partial sequences of G9 aptamers which showed the highest binding signals among the population. However those conserved regions did not show

significant binding to *S. mutans* (Table S3). Thus, it was suggested that those guanine rich sequences did not bind to *S. mutans* alone and variable termini were also important for binding ability of aptamers. Furthermore, the sequence of SMa#G9-87 which was the highest binding aptamer contained a conserved guanine rich sequence (5´-GGGGGGAGGGGGGG-3´) flanked by variable sequences, and this certain sequence appeared in other G9 aptamers with relatively high frequency. It was also found in SMa#3-6-P1, one of Cell-SELEX selected parent aptamers. SMa#G9-87 showed a 4.5-fold increase in binding signal to *S. mutans* compared with SMa#3-6-P1 **(**Fig. 2A, Table S3). In addition, it is known that those consecutive guanine blocks potentially form guanine quadruplex (Gq) structures, which can be potential scaffolds of aptamer structures (Ikebukuro et al. 2005; Yoshida et al. 2013). For these reasons, we considered this sequence (5´-GGGGGGAGGGGGGG-3´) would be a core sequence, which is important sequence to determine aptamer scaffold for target recognition, and other sequence parts would be also important for aptamer binding. Thus, we produced the subsequent generation of ISM to optimize sequence of variable termini by ISM.

We performed further ISM to optimize aptamer sequence by another shuffling method in the G10 to G12. This time we shuffled each variable region while fixing the core sequence to adapt variable regions for specific binding to *S. mutans*. Each top five of all sequences with highest affinity and specificity were listed in Table 1. As a result of 11 cycles of ISM, two G11 aptamers, SMa#G11-5 and SMa#G11-6, showed the highest binding signals to *S. mutans* with a 16-fold increase over a Cell-SELEX-selected aptamer SMa#3-6-P1 which also has the core sequence. This represents even a 10-fold increase over the top aptamer of parents (Fig. 2A, Table S2). The G11 aptamers displayed an up to 12-fold more binding to *S. mutans* than to *L. acidophilus* (Fig. 2B), representing a 1.5-fold increase of specificity over parent aptamers. Those findings indicated the ISM successfully improved aptamer affinity for *S. mutans* and specificity at the same time, and we identified aptamers showing significantly higher affinity and specificity compared with Cell-SELEX-selected aptamers (Fig. 2C).

Moreover, lengths of all of the G11 sequences were become 44- or 47-nt (Table S3), although ISM was initiated from parents with a range of lengths between 34- to 80-nt (Table S2). This finding suggests the process of ISM optimized length of aptamers as well as sequences to show improved affinity and specificity. It is remarkable that ISM based on GAs would be a powerful approach to optimize aptamer length at the same time of sequence optimization, which is usually difficult to achieve by conventional techniques such as site-directed mutagenesis.

Because G12 sequences did not show improvement of binding ability nor specificity, we accordingly concluded that ISM achieved a large space searching of oligonucleotide sequences from G1 to G6 to increase sequence diversity, and the convergence and optimization of aptamer sequences with the core sequence by G11. During ISM, we set both affinity and specificity as fitness and successfully improved them by generating and evaluating a total of 300 aptamer sequences **(**Fig. 2C, Table S3). The discovery of the core sequence accelerated the improvement of the aptamers. This type of sequence convergence is well known as a schema representing a mile stone during GA operation (Goldberg 1989; Holland 1975). The discovery of a schema is the key factor in a GA, which positively affects evolution of high fitness populations; therefore, our finding appears reasonable from this perspective.

Characterization of improved aptamers

To characterize the binding affinities of SMa#G11-5 and SMa#G11-6, we performed a flow cytometry analysis obtaining a dissociation constant (K_d) value of 33 nM for both aptamers (Fig. S1). Binding specificity of these aptamers was also assessed by aptamer blotting assays. Both aptamers displayed binding to all tested strains of *S. mutans* but not to other species, whereas *L. acidophilus* still showed binding to these aptamers with a low binding signal (Fig. 3). These results indicate that evolved aptamers are specific to *S. mutans* cells, but *L. acidophilus* cells may be inclined to bind

nonspecifically to DNAs. This inference is in agreement with the result of an aptamer blotting assay indicating binding of an ssDNA library to *L. acidophilus* (Fig. S2).

To elucidate the structural conformation of the improved aptamers, we acquired circular dichroism (CD) spectra of SMa#G11-5 and SMa#G11-6. These spectra suggested the formation of parallel-type Gq structures (Paramasivan et al. 2007) (Fig. S3A, B). We further characterized the conformation of the aptamers by native polyacrylamide gel (PAGE) analysis using a BODIPY-labeled macrocyclic heptaoxazole, which is a fluorescence probe that interacts specifically with Gq structures (Tera et al. 2010). When aptamers were incubated with the probe, broad fluorescence bands were observed by PAGE analysis (Fig. S4A, B), suggesting the formation of polymerized Gq structures. This finding indicated SMa#G11-5 and SMa#G11-6 form polymerized parallel Gq structures. Although the core sequence itself was found to form similar structures, it did not display appreciable binding to *S. mutans* **(**Fig. S3C, Fig. S4A, B). These suggested the variable termini of the aptamers are required for specific binding to *S. mutans* and the core sequence of an aptamer determines its three-dimensional conformation. It is known that a polymerized Gq structure can change its nano-structure depending on sequence and structure of other parts of the sequence (Morita et al. 2011). We considered that the appropriate combination of the core sequence and its variable termini was necessary for folding into a binding structure of aptamers showing high affinity and specificity for *S. mutans*. Therefore, we believe ISM successfully identified such optimal sequence through successive rounds of GA operations. We have suggested aptamers with Gq-structures tend to preferentially bind to β-structures of proteins (Tsukakoshi et al. 2012; Yoshida et al. 2013). Thus, the aptamers identified in this study may interact with a β-structure of membrane associated protein of *S. mutans*, although it requires further studies to identify the target molecule of aptamer binding.

Detection of *S. mutans* using aptamer immobilized beads and gold colloid

We constructed a biosensing platform for detection of *S. mutans* using the evolved aptamers. First, we demonstrated the microscopic detection of *S. mutans* using 5′-biotin-modified SMa#G6-25 immobilized on avidin agarose beads to confirm the possibility of selective detection of *S. mutans* over *L. acidophilus*. The aptamer-immobilized beads selectively captured *S. mutans* (Fig. 4A, B). This indicated that aptamers with improved affinity and specificity have potential for specific detection of *S. mutans*. We accordingly constructed a flow-through detection platform using the evolved aptamer SMa#G11-5 immobilized on gold colloid. *S. mutans* incubated with gold-colloid-labeled SMa#G11-5 were applied to flow-through tests using a membrane with pore size of 0.45 μm able to capture *S. mutans* cells. We performed colorimetric detection of gold colloids that bound to target cells and remained on the membrane (Fig. 5A). The assay demonstrated detection of *S. mutans* cells at a concentration range of 1×10^5–10^8 CFU/mL (Fig. 5B). The colorimetric assays using gold-colloid-labeled aptamers do not require sophisticated detection apparatus (Balamurugan et al. 2008), and the detection can be performed in solution, avoiding the disadvantages of platform immobilization. This aptamer based flow-through test would ultimately be an effective sensing platform for detection of *S. mutans*.

Conclusions

In conclusion, we succeeded in improvement of affinity and specificity of aptamers against *S. mutans* using ISM. This is the primary study of the development of aptamers recognizing *S. mutans* cells and also the simultaneous improvement of binding specificity and affinity of aptamers by a directed evolutionary approach. *In silico* sequence modifications facilitated a large space searching of aptamer sequences to yield a sequence group that favors for specific binding. The discovery of a core sequence accelerated the improvement process. This is the first report of aptamer improvement based on a core

sequence discovered during *in silico* sequence modification. The use of an evolution-mimicking algorithm in ISM that fixes specific sequence blocks facilitated the convergence of aptamer sequences with a specific structural topology and optimized aptamer sequences for target binding. Moreover, we demonstrated specific detection of *S. mutans* by flow-through tests at a concentration range of 1×10^5–10^8 CFU/mL using the evolved aptamer. These results indicate the potential of ISM for aptamer development for biosensing applications. We believe that the ISM will advance selection of aptamers with functions of interest such as high specificity and affinity with parental sequences selected by SELEX or other techniques.

Acknowledgements

We thank Prof. K. Nagasawa (Tokyo University of Agriculture and Technology) for providing the Gq ligand. This study was supported by the Industrial Technology Research Grant Project 2009 from the New Energy and Industrial Technology Development Organization of Japan (NEDO) and a grant from the Low-Carbon Research Network Japan (LCnet) for K. I. N.S. was supported by the Japan Society for Promotion of Science (JSPS) Research Fellowships for Young Scientists DC1.

References

Ahmad KM, Xiao Y, Soh HT. 2012. Selection is more intelligent than design: improving the affinity of a bivalent ligand through directed evolution. Nucleic Acids Res 40(22):11777-83.

Balamurugan S, Obubuafo A, Soper SA, Spivak DA. 2008. Surface immobilization methods for aptamer diagnostic applications. Anal Bioanal Chem 390(4):1009-21.

Chen F, Zhou J, Luo F, Mohammed AB, Zhang XL. 2007. Aptamer from whole-bacterium SELEX as new therapeutic reagent against virulent Mycobacterium tuberculosis. Biochem Biophys Res Commun 357(3):743-8.

Cho M, Xiao Y, Nie J, Stewart R, Csordas AT, Oh SS, Thomson JA, Soh HT. 2010. Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. Proc Natl Acad Sci U S A 107(35):15373-8.

Ellington AD, Szostak JW. 1990. In vitro selection of RNA molecules that bind specific ligands. Nature 346(6287):818-22.

Fang X, Tan W. 2010. Aptamers generated from cell-SELEX for molecular medicine: a chemical biology approach. Acc Chem Res 43(1):48-57.

Gevertz J, Gan HH, Schlick T. 2005. In vitro RNA random pools are not structurally diverse: a computational analysis. RNA 11(6):853-63.

Goldberg DE. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Boston: Addison-Wesley Longman Publishing Co., Inc. .

Hamada S, Slade HD. 1980. Biology, immunology, and cariogenicity of Streptococcus mutans. Microbiol Rev 44(2):331-84.

Holland JH. 1975. Adaptation in natural and artificial systems. Ann Arbor: University of Michigan Press.

Ikebukuro K, Okumura Y, Sumikura K, Karube I. 2005. A novel method of screening thrombin-inhibiting DNA aptamers using an evolution-mimicking algorithm. Nucleic Acids Res 33(12):e108.

Knight CG, Platt M, Rowe W, Wedge DC, Khan F, Day PJ, McShea A, Knowles J, Kell DB. 2009. Array-based evolution of DNA aptamers allows modelling of an explicit sequence-fitness landscape. Nucleic Acids Res 37(1):e6.

Kunii T, Ogura S, Mie M, Kobatake E. 2011. Selection of DNA aptamers recognizing small cell lung cancer using living cell-SELEX. Analyst 136(7):1310-2.

Mayer G, Ahmed MS, Dolf A, Endl E, Knolle PA, Famulok M. 2010. Fluorescence-activated cell sorting for aptamer SELEX with cell mixtures. Nat Protoc 5(12):1993-2004.

Morita Y, Yoshida W, Savory N, Han SW, Tera M, Nagasawa K, Nakamura C, Sode K, Ikebukuro K. 2011. Development of a novel biosensing system based on the structural change of a polymerized guanine-quadruplex DNA nanostructure. Biosens Bioelectron 26(12):4837–41.

Noma T, Ikebukuro K. 2006. Aptamer selection based on inhibitory activity using an evolution-mimicking algorithm. Biochem Biophys Res Commun 347(1):226-31.

Noma T, Sode K, Ikebukuro K. 2006. Characterization and application of aptamers for Taq DNA polymerase selected using an evolution-mimicking algorithm. Biotechnol Lett 28(23):1939-44.

Nonaka Y, Yoshida W, Abe K, Ferri S, Schulze H, Bachmann TT, Ikebukuro K. 2013. Affinity improvement of a VEGF aptamer by in silico maturation for a sensitive VEGF-detection system. Anal Chem 85(2):1132–7.

Ogasawara D, Hasegawa H, Kaneko K, Sode K, Ikebukuro K. 2007. Screening of DNA aptamer against mouse prion protein by competitive selection. Prion 1(4):248-54.

Paramasivan S, Rujan I, Bolton PH. 2007. Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. Methods 43(4):324-31.

Platt M, Rowe W, Wedge DC, Kell DB, Knowles J, Day PJ. 2009. Aptamer evolution for array-based diagnostics. Anal Biochem 390(2):203-5.

Savory N, Abe K, Sode K, Ikebukuro K. 2010. Selection of DNA aptamer against prostate specific antigen using a genetic algorithm and application to sensing. Biosens Bioelectron 26(4):1386-91.

Savory N, Lednor D, Tsukakoshi K, Abe K, Yoshida W, Ferri S, Jones BV, Ikebukuro K. 2013. In Silico Maturation of Binding-Specificity of DNA Aptamers Against Proteus mirabilis. Biotechnol Bioeng. doi: 10.1002/bit.24922

Schutze T, Wilhelm B, Greiner N, Braun H, Peter F, Morl M, Erdmann VA, Lehrach H, Konthur Z, Menger M and others. 2011. Probing the SELEX process with next-generation sequencing. PLoS One 6(12):e29604.

Shangguan D, Li Y, Tang Z, Cao ZC, Chen HW, Mallikaratchy P, Sefah K, Yang CJ, Tan W. 2006. Aptamers evolved from live cells as effective molecular probes for cancer study. Proc Natl Acad Sci U S A 103(32):11838-43.

Tera M, Iida K, Ikebukuro K, Seimiya H, Shin-Ya K, Nagasawa K. 2010. Visualization of G-quadruplexes by using a BODIPY-labeled macrocyclic heptaoxazole. Org Biomol Chem 8(12):2749-55.

Torres-Chavolla E, Alocilja EC. 2009. Aptasensors for detection of microbial and viral pathogens. Biosens Bioelectron 24(11):3175-82.

Tsukakoshi K, Abe K, Sode K, Ikebukuro K. 2012. Selection of DNA aptamers that recognize α-synuclein oligomers using a competitive screening method. Anal Chem 84(13):5542–7.

Tuerk C, Gold L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249(4968):505-10.

Yoshida W, Saito T, Yokoyama T, Ferri S, Ikebukuro K. 2013. Aptamer selection based on g4-forming promoter region. PLoS One 8(6):e65497.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406–15.

Table and Figure legends

Table 1. Each top-five aptamer sequences showing highest binding ability and specificity for *S. mutans*.

Fig. 1. Schematic of *in silico* sequence modification that produces each generation of sequences in *in silico* maturation. The G1–G6 generation sequences were produced by crossover and point mutations to explore a wide sequence space, G7–G9 were produced by shuffling with fixing conserved sequence blocks to optimize aptamer sequences and to discover a core sequence, and G10–G12 were produced by shuffling with the core sequence fixed to optimize variable regions for the discovered core sequence. All sequences in this figure are just schematic representations.

Fig. 2. *In silico* maturation of DNA aptamers for improving binding affinity and specificity for *S. mutans.* Evolution of (A) binding ability to *S. mutans* and (B) specificity over *L. acidophilus*. (C) Correlation of binding ability and specificity of parent aptamers (△), G1–G10 sequences (○) and G11 sequences (●). Binding ability to *S. mutans* and *L. acidophilus* of each sequence was evaluated by aptamer blotting assays. Relative binding ability was normalized by binding signal of a parent aptamer SMa#3-6-P1 as 1.00. Relative binding specificity represents a ratio of binding signal for *S. mutans* to that for *L. acidophilus* of each aptamer.

Fig. 3. Aptamer blotting assay for binding of SMa#G11-5 and SMa#G11-6 to *S. mutans* and other *Streptococcus* species and *L. acidophilus* cells. 1×10^{10} CFU of each bacterial species were spotted on

nitrocellulose membrane. After incubation with 5′- fluorescein-modified SMa#G5-11 or SMa#11-6 (f.c. 50 nM), aptamer binding was detected using anti-FITC antibody HRP conjugate.

Fig. 4. Microscopic assay for binding of SMa#G6-25-immobilized agarose avidin beads to (A) *S. mutans* and (B) *L. acidophilus*. 5′-Biotin-modified SMa#G6-25 was immobilized on avidin agarose beads and incubated with *S. mutans* or *L. acidophilus*. After washing, cells captured on the beads were analyzed by microscopy at a magnification of 1000×.

Fig. 5. Detection of *S. mutans* by flow-through test using SMa#G11-5 modified gold colloid. (A) 5′-Thiol-modified SMa#G11-5 was immobilized on colloidal gold and incubated with *S. mutans* cells following a flow-through test with the PES membrane. (B) Spot intensity was assessed using ImageQuant TL software and plotted against *S. mutans* concentration. Error bars show standard deviations (n = 3).

Table 1. Each top-five aptamer sequences showing highest binding ability and specificity for *S. mutans*.

	Name	Generation	Relative binding ability (A.U.)	Relative binding specificity (A.U)	Sequence
Top 5 (Binding ability)	**SMa#G11-6**	11	16.4	12.3	ATACTATCGCATTCCTTCCG AGGGGGGAGGGGGGGGTG GGGGTCGGT
	SMa#G11-5	11	14.0	10.7	ATACATCTTAAGTCTCGTGG GGGGAGGGGGGGTTGGTG GGCTTT
	SMa#G11-20	11	8.86	5.29	CTACGTCTAGATTCCAGTCG AGGGGGGAGGGGGGGTTT TGGATCGGT
	SMa#G10-6	10	8.39	6.72	ACACCAGCGTATTCTCTTG GGGGGAGGGGGGGTTGGG GGTCGGT
	SMa#G10-11	10	7.03	2.62	CAACCTGCTTATCTAAGGG GGGGAGGGGGGGTTGTGG GTAGGT
Top 5 (Binding specificity)	**SMa#G11-6**	11	16.4	12.3	ATACTATCGCATTCCTTCCG AGGGGGGAGGGGGGGGTG GGGGTCGGT
	SMa#G2-1	2	0.96	11.5	CTAGCAGTTCATTCAATAGG GGGGACGGGGGGTTAGGCT AGACTATGTCTAATCA
	SMa#G11-5	11	14.0	10.7	ATACATCTTAAGTCTCGTGG GGGGAGGGGGGGTTGGTG GGCTTT
	SMa#G5-5	5	1.59	10.2	ATACACGTAAGTACCATTG GGCGGGGGGGGTGTTGTTT TCTACTATGTGCAATCG
	SMa#G11-8	11	6.94	8.26	ACACCACTTAATTCCATGC GAGGGGGGAGGGGGGGTT CGGGGACGGC

Binding ability and specificity were assessed by aptamer blotting assays. Relative binding ability was normalized by binding signal of a parent aptamer SMa#3-6-P1 as 1.00. Relative binding specificity is a ratio of binding signal for *S. mutans* to that for *L. acidophilus* of each aptamer.

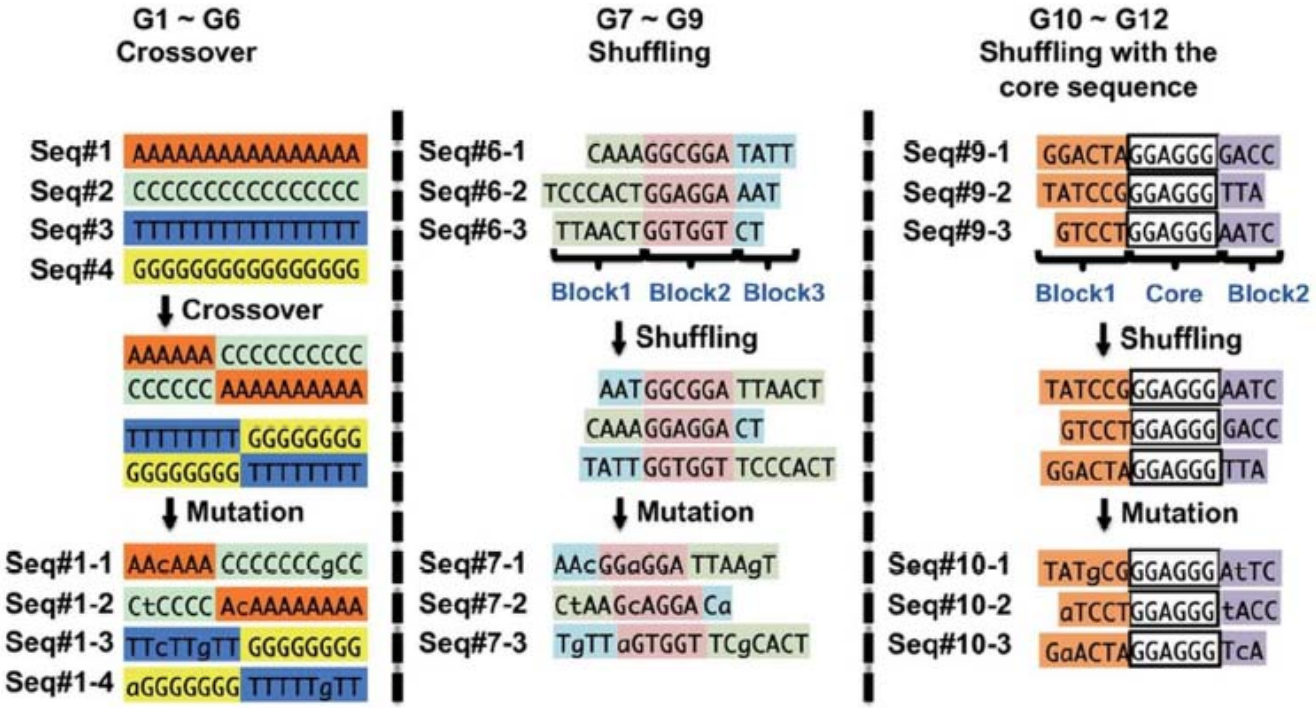

Figure 1

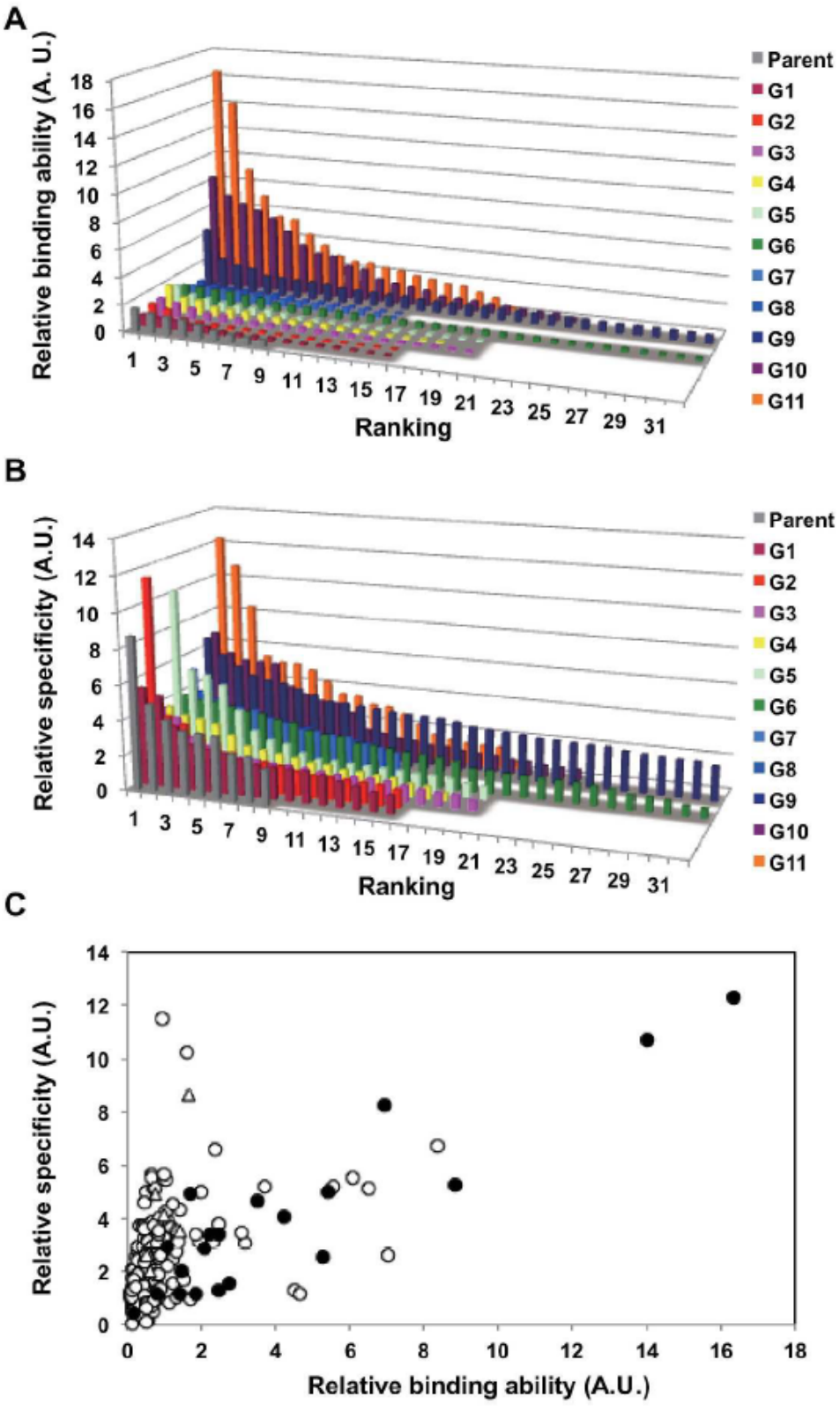

Figure 2

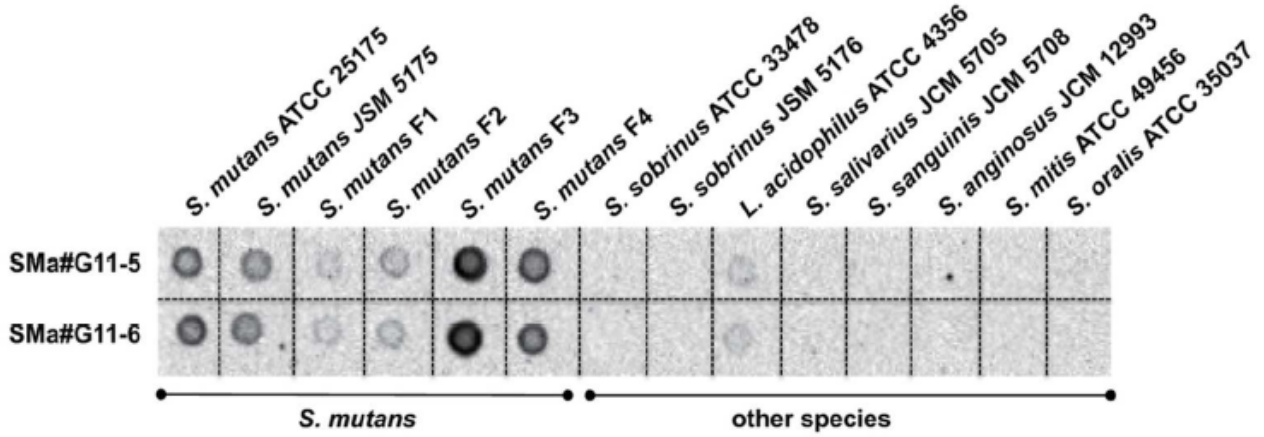

Figure 3

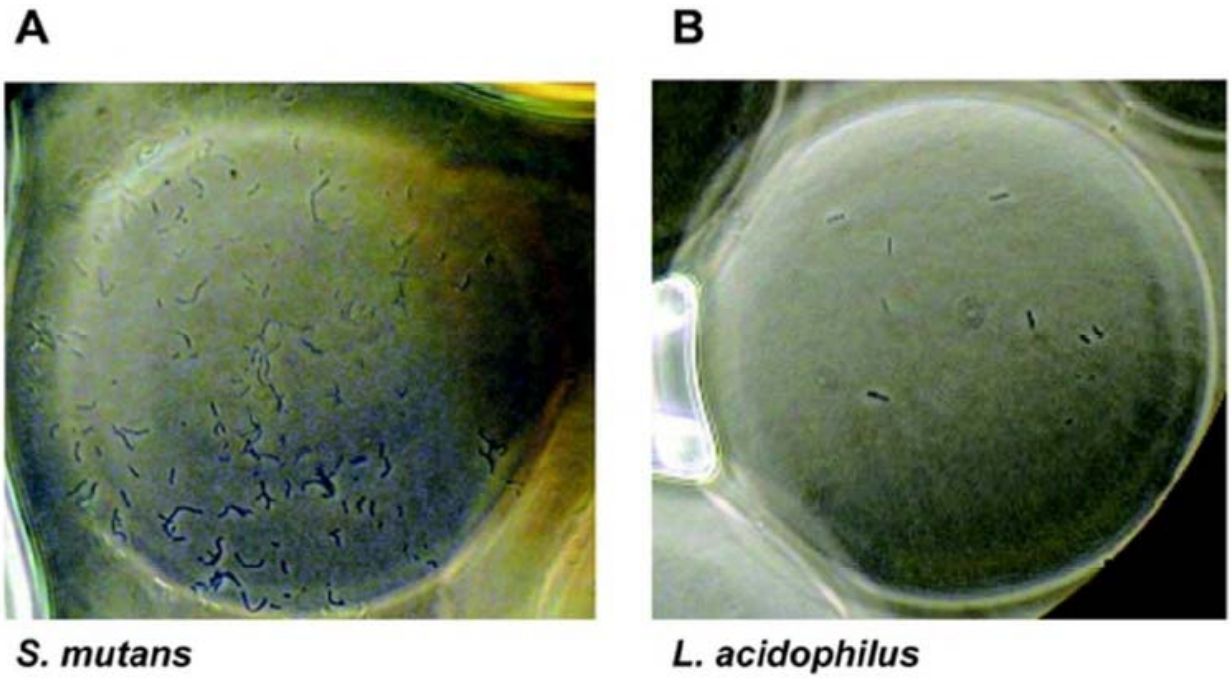

Figure 4

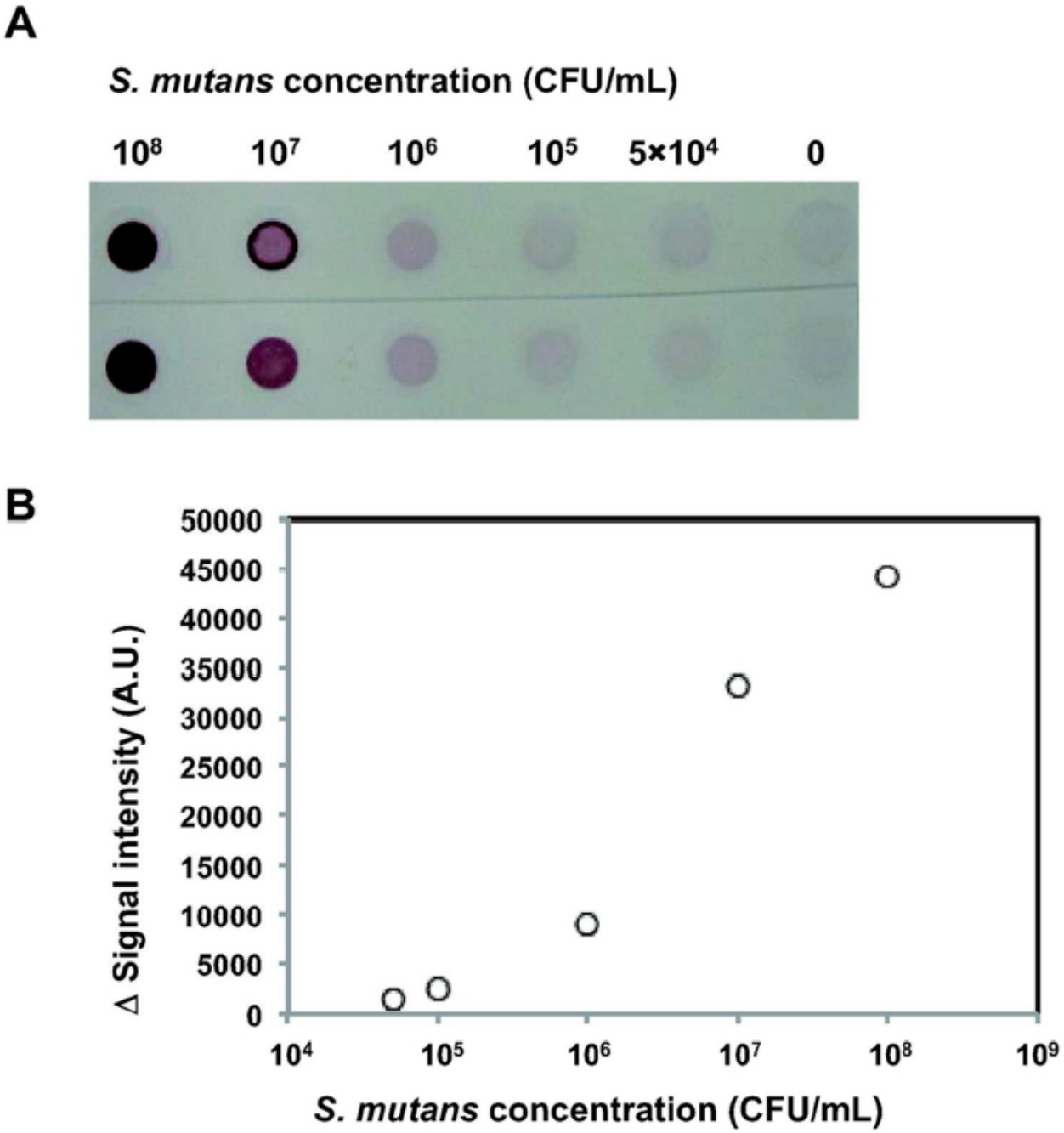

Figure 5