

ゲノム配列に基づく楽曲生成技術の開発  
~音楽によって表される生物の進化~

環境情報学部 4年

高橋佑歌

## 1. 研究の背景

近年次世代シーケンサーの発達によって公共データベースに掲載されている生物の配列数は急激に増加しており、これにより既存の配列情報を使った情報学的なアプローチによるゲノム解析や、多数の配列を用いた比較系統解析が可能になった。その一方で、得られた生命情報の量に対して解決すべき課題は依然として多く残っているのが現状である。ここで注目したいのは、配列解析のアウトプット方法が視覚情報に集中している点である。

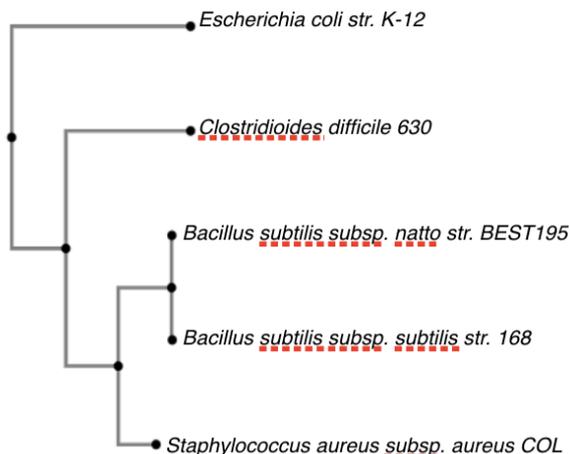
私は視覚情報に代わるゲノム情報の新しい表現方法として、聴覚情報を提案する。塩基配列を一定のアルゴリズムに従って音の情報に変換し、ゲノム配列に沿った音楽を生成するのである。塩基配列は4つの文字からなる文字列であり、実際に生体内では一定のルールに従ってそれらの情報を元に約20種類のアミノ酸が生産される。それぞれのアミノ酸に対応した3塩基のコドンが複数存在し、読み枠に沿ってそれらの情報が読み取られるのだ。配列情報からアミノ酸への変換を行っているように、塩基配列からの変換の方法を規定するルールを設けることで、音楽を生成することは可能である。音楽を用いることで、似たような配列の繰り返しなど、それぞれゲノムの特徴を表現することができるため、目的の配列に特異的な性質を見つける上で有用であると考ええる。また、ゲノム配列を音楽に変換することで、ゲノム配列の特徴である繰り返しのパターンが認知しやすくなると考える。目視で1遺伝子レベルでも数百塩基以上であるゲノム配列を追うことは極めて困難であり、音楽を用いることによってこういったゲノムの特性をシンプルに、分かりやすく認知することができると考えられる。このことから、ゲノム音楽は研究分野でのアウトプット手法としてだけでなく、科学に興味を持つきっかけとしても初等科学教育にも応用が期待できる。先行研究としてはヒトゲノムから音楽を作成した例などが報告されている [1],

本研究では、単一の配列から作成する既存のゲノム音楽とは異なり、複数の生物から得られる同一遺伝子配列から1つの音楽を作成することで、進化の様相を音楽で表すことを目標とする。生物のゲノムは時間経過とともに変異が蓄積するため、一般的に系統的に遠い生物種であるほど配列の類似性が低くなる。そこで、今回は普遍的な遺伝子である DNA polymerase I の配列を用いて音楽で表現することで、遺伝子の進化の様相を表現することを試みた。

## 2. 結果

### 2.1 系統樹の作成

米国生物科学情報センター(NCBI)の GenBank データベースからダウンロードした6種の生物ゲノム(*Bacillus subtilis subsp. natto str. BEST195*, *Bacillus subtilis subsp. subtilis str. 168*, *Staphylococcus aureus subsp. aureus COL*, *Clostridioides difficile 630*, *Escherichia coli str. K-12*)を解析の対象とした。これらの16S rRNA配列を用いて最尤法で系統樹を作成した結果を以下に示す。



## 2.2 ゲノム音楽の作成

ゲノム配列から DNA polymerase I の配列を抽出し、以下の表に従って配列を音に変換し、それらを結合することで楽曲を作成した。今年度は予算の都合上 MIDI を用いての楽曲編集には至らなかったため、これらの情報は ABC 譜に書き起こした。紙面の都合上譜面は以下の URL に掲載する。

音	アミノ酸	コドン
C	Asp, GLu	GAU, GAC, GAA, GAG
C #	Cys, Met	UAU, UGC, UGC AUG
D	Lys, Arg	AAA, AAG, AGA, AGG, CGU, CGC, CGA, CGG
D #	Ser, Thr	AGU, AGC, ACU, ACC, ACA, ACG, UCU, UCC, UCA, UCG
E	Tyr, Phe	UAU, UAC, UUU, UUC
F	Asn, Gln	AAU, AAC, CAA, CAG
F #	Gly	GGU, GGC, GGA, GGG
G	Leu, Val, Ile	CUU, CUC, CUA, CUG, GUU, GUC, GUA, GUG, AUU, AUC, AUA, UUA, UUG
G #	Ala	GCU, GCC, GCA, GCG
A	Trp, 停止	UGG, UAA, UAG, UGA
B ♭	His	CAU, CAC
B	Pro	CCU, CCC, CCA, CCG

譜面 URL [https://drive.google.com/open?id=1AFJvncRDDJCC\\_RN9afMe6Vy5lg\\_s8JJn](https://drive.google.com/open?id=1AFJvncRDDJCC_RN9afMe6Vy5lg_s8JJn)

### 参考文献

- [1] Gena, Peter, and Charles Strom. "Musical synthesis of DNA sequences." *XI Colloquio di Informatica Musicale* (1995): 203-204.