Universitetet i Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY
MASTER'S THESIS

Study programme/specialization: **Biological Chemistry**	Spring semester **2019** **Open**
Author: **Inger Djuve Aanderaa**	..... (signature of author)
Faculty supervisor: **Svein Bjelland**	
Title of master's thesis: **Molecular biology of modified DNA bases**	
Credits: **60**	
Keywords: **f^{5}dU induced mutagenesis** **Nucleotide excision repair** **Base excision repair** **5-formyluracil** **MutY** **hMPG** **hSMUG (25-270)** ***N*^{4},5-dimethylcytosine**	Pages: **54** + supplementary materials: **208** **Stavanger, july 2019**

Acknowledgement

The outcome of this master thesis is the result of extensive lab work performed at the Centre for Organelle Research (CORE) that belongs to the Department of Mathematics and Natural Sciences, Faculty of Science and Technology, University of Stavanger, Norway.

I would first like to express my deep gratitude to our research group leader, and my main supervisor Professor Svein Bjelland, for once again trusting me to be part of his team. Even though our group is small, he provides us with the opportunity and support to explore the interesting topics of DNA damage and repair. He has been a great advisor regarding all theoretical topics, and even though we don't see much of him in the lab, he takes a great interest in our work, progression and results. I would also like to thank him for his patience with my never ending array of questions regarding mutation induction, DNA repair and everything even remotely related to these topics, from a student who will probably never learn to tame her curious mind to stay on topic. I have also greatly appreciated our long talks and discussions about politics and other general topics.

I would also like to thank my supervisors in the lab, Marina Alexeeva, PhD and Almaz Tesfahun, PhD student, for all guidance regarding protein production and purification, and DNA incision/excision assays, respectively. I would also like to thank them for all assistance and comfort through challenging times in the lab when things did not go according to plan, and finally for long talks about life outside the lab. Even though our group is small it is part of a larger collaboration at CORE, I would therefore like to thank all engineers and technicians for technical assistance, encouraging words and clean and sterilized equipment.

This thesis would not be possible without the support of my family, and therefore I share the achievement with them. My wife Karina has been a great motivator and her ability to reset my mind when it gets stuck is invaluable to me. I am endlessly indebted to her for understanding my absence during evenings and weekends, spent either in the laboratory or reading and writing in my home office. To my children Sigmund Arne and Sigrid Helene, thank you for your hugs and smiles, which can drive away any grievances, and thank you for reminding me what is truly important in life.

I also have to thank my mother, brother and sister for always supporting me, believing in me and pushing me to do my very best, and my in law family for encouraging words, support and much needed babysitting.

Abstract

The cell has several mechanisms for repairing DNA damages, and it has been widely accepted that the base excision repair (BER) pathway is the main repair mechanism for small non helix distorting damages caused by oxidation. However, evidence exists that the nucleotide excision repair (NER) pathway plays a role in repair of oxidized bases in mammalian cells *in vitro*. To investigate if the NER pathway is involved in repair of oxidized bases *in viv*o, exponentially growing *E. coli* cells were exposed to 0.1 mM 5-formyldeoxyuridine (f^5dU) followed by analyzing the mutations caused in the *rpoB*$^+$ gene. Previous experiments performed within our group indicated that the *uvrA*$^+$ gene was highly involved in promoting mutations caused by exposure to 5-formyldeoxyuridine. This thesis work combined with results from previously performed experiments within the same study, show that all the three Uvr proteins are involved in promoting AT→GC mutations at sites 1534 and 1547 within the *rpoB*$^+$ gene following exposure to a low concentration of f^5dU. An indication that there might also be an unknown function of UvrA is something that needs to be further investigated.

5-methylcytosin (m^5C) is recognized as the most important epigenetic DNA base in mammalian cells, yet relatively few studies have investigated damaging chemical alterations to this important DNA base. Certain methylases expressed by prokaryotes can convert m^5C to N^4,5-dimethylcytosine ($m^{N4,5}C$) *in vitro* and it is therefore a possibility of $m^{N4,5}C$ existing *in vivo*. MutY, hMPG and a truncated version of hSMUG1 (hSMUG 25-270) were investigated for activity against $m^{N4,5}C$ *in vitro.* Short fluorescently tagged oligodeoxyribonucleotides with $m^{N4,5}C$ inserted at a specific position were hybridized with complimentary oligodeoxyribonucleotides with A, C, G or T placed opposite the damaged base, followed by incubation with the abovementioned enzymes. Denaturing PAGE was used to determine incision activity. None of the tested enzymes showed any activity against the lesion.

List of figures

List of tables

Abbreviations

5'-dRP – 5'-deoxy ribose phosphate
5-OH-dHT – 5-hydroxy-5,6-dihydrothymine
5-OH-dHU – 5-hydroxy-5,6-dihydrouracil
6-OH-dHC – 6-hydroxy-5,6-dihydrocytosine
6-OH-dHT – 6-hydroxy-5,6-dihydrothymine
6-OH-dHU – 6-hydroxy-5,6-dihydrouracil
εA – 1,*N*6-ethenoadenine
εC – 3,*N*4-ethenocytosine
AlkA - 3-methyladenine glycosylase II
AP sites – Apurinic- or Apyrimidinic- (Abasic) sites in DNA
APE – AP endonuclease
ATP – Adenosine triphosphate
BER – Base Excision Repair
Ca^5C – 5-carboxymethylcytosine
Cg – cytosine glycol
CPD - Cyclobutane Pyrimidine Dimers
dhC – 5,6-dihydroxycytosine
dHT – 5,6-dihydrothymine
dHU – 5,6-dihydrouracil
ds – double stranded
f^5C – 5-formylcytosine
f^5dU – 5-formyl-2'-deoxyuridine
f^5U – 5-formyluracil
FapyA – 4,6-diamino-5-formidopyrimidine
FapyG – 2,6-diamino-4-hydroxy-5-formamidopyrimidine
Fpg - formamidopyrimidine-DNA glycosylase
GG-NER – Global Genomic Nucleotide Excision Repair
Gh – guanidinohydantoin
h^5C – 5-hydroxycytosine
h^5U - 5-hydroxyuracil
hm^5U – 5-hydroxymethyluracil
hmh – 5-hydroxy-5-methylhydantoin
HtH – Helix-turn-Helix
m^5C – 5-methylcytosine
h^5C – 5-hydroxycytosine
hm^5U - 5-hydroxymethyluracil
m^3A – 3-methyladenine
m^3G – 3-methylguanine
m^7G – 7-methylguanine
m^7A – 7-methyladenine
mfapy – 2,6-diamino-4-oxo-5-(N-methyl)formamidopyrimidine
$m^{O2}C$ – *O*2-methylcytosine
$m^{O2}G$ – *O*2-methylguanine
$m^{O4}T$ – *O*4-methylthymine
$m^{N6}A$ – 6-methyladenine
Nei – Endonuclease VIII
NER – Nucleotide Excision Repair
Nth – Endonuclease III
Oxo^8A – 7,8-dihydro-8-oxoadenine
Oxo^8G – 7,8-dihydro-8-oxoguanine
PAH - Polycyclic aromatic hydrocarbons
Pol I – DNA Polymerase I
Rif^R – Rifampicin resistance
ROS – Reactive Oxygen Species
SAM – S-adenosylmethionine
Sp – spiroiminodihydantoin
Ss – single stranded
Tag - 3-methyladenine glycosylase I
TC-NER – Transcription Coupled Nucleotide Excision Repair
Th^5 – 5-hydroksy-6-hydrothymine
Tg – Thymine Glycol
UDG – Uracil-DNA Glycosylase
Ug – 5,6-dihydroxy-5,6-dihydrouracil (uracil glycol)
Uh^5 – 5-hydroxy-6-hydrouracil
UV – Ultra Violet
UvrD – DNA Helicase II

1 Introduction

1.1 DNA Damages

DNA damages are erroneous alterations in the chemical structure of DNA. Important lesions are single- and double-strand breaks, apurinic or apyrimidinic (AP) sites, and chemical alteration to bases. If the damage is not repaired prior to replication, it may cause arrest of DNA replication or transcription resulting in cell death (classified as cytotoxic damage), or mutagenesis (classified as mutagenic damage). The latter can also have devastating consequences for the cell itself, or for the entire organism harboring that cell. DNA damages can be caused by both exogenous and endogenous agents. The exogenous agents and their effects have been extensively studied as they have major effects and are in many cases easily preventable. Such agents include ionizing radiation and various carcinogens. Ionizing radiation can cause oxidative damages, DNA base dimers, single strand breaks and double strand breaks, the latter regarded as the most devastating damage. Cigarette smoke contains various carcinogens including the extensively studied polycyclic aromatic hydrocarbons (PAH's), which can attach to DNA and form adducts. As a defense mechanism PAH's are hydroxylated by Cytochrome P450 enzymes to make them more water soluble, but this may result in other harmful intermediates (e.g. benzo(α)pyrene diol epoxide) efficiently creating adducts in DNA. Studies conducted on various exogenous agents have resulted in e.g. restrictions in cigarette sales (and use), a ban on formerly widely used compounds such as asbestos, and informational campaigns regarding preventive measures and limitation to sun exposure. Whereas measures can be taken to limit or prevent the exposure to most exogenous agents, the threat posed by endogenous agents (i.e., being a consequence of the cell chemistry itself), cannot be avoided. Such damages may arise from hydrolytic reactions (principally, the cell is an aqueous gel), alkylation- or methylation-reactions and oxidation reactions. Hydrolytic reactions cleave chemical bonds in DNA, leading to deamination products (e.g. deamination of cytosine forms uracil) or loss of individual bases (N-glycosidic bond cleavage leads to an AP site). Alkylation- or methylation- reactions can alter bases and give rise to mutations, as well as induce AP sites.

1.1.1 Oxidized bases

Reactive oxygen species (ROS) can attack all the normal bases in DNA leading to a variety of oxidative damages. The most important source of ROS is aerobic respiration, which in eukaryotes take place in the mithochondria. Thus, some electrons are transferred directly to O_2 creating the superoxide ion O_2^-, which is then converted to H_2O_2 through the superoxide dismutation reaction, the portion of H_2O_2 which is not removed enzymatically can combine with Fe^{2+} through the Fenton reaction and cause the hydroxyl radical •OH ($O_2^- \rightarrow H_2O_2 \rightarrow$ •OH). Exposure of water, and thus cells, to ionizing radiation will also create •OH, which is the most reactive chemical species known, able to react with numerous biomolecules, including DNA and its precursors. The different ROS-oxidized bases pose different challenges to the genomic integrity. A small selection of the most studied are shown in Figure 1.

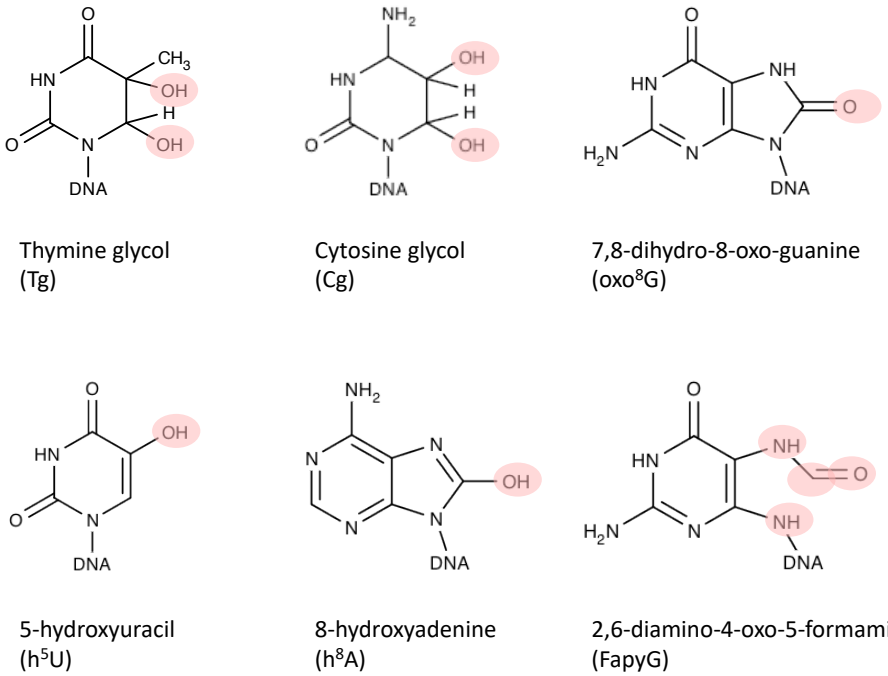


Figure 1: Common oxidized bases
Red circle indicates the modified position

Thymine is oxidized at two distinct sites; the 5,6-double bond, and the 5-methyl group. Abundant damages arising from these attacks are among others: thymine glycol (Tg), 5,6-dihydrothymine (dHT), 5-hydroxy-5,6-dihydrothymine (5-OH-dHT), 6-hydroxy-5,6-dihydrothymine (6-OH-dHT) and 5-hydroxy-5-methylhydantoin (hmh) [1]. The absence of the methyl group renders the 5,6-double bond as the major site of oxidative attack on cytosine, which changes the planar aromatic ring structure into a non-planar non-aromatic structure that leads to the 5,6-dihydroxy-5,6dihydrouracil (cytosine glycol, Cg) lesion which can be further deaminated and dehydrated to 5,6-dihydroxy-5,6-dihydrouracil (uracil glycol, Ug) and 5-hydroxycytosine (h^5U), respectively [1]. The epigenetic mark 5-methylcytosine (m^5C) is an interesting hybrid of thymine and cytosine by exhibiting both targets of attack. In guanine ROS react with C8 resulting in a C8-OH adduct radical, which can be further oxidized or reduced to 7,8-dihydro-8-oxodeoxyguanine (oxo^8G) or 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG), respectively, the two major oxidative damages of guanine [1]. Adenine oxidation is similar to guanine oxidation and yields the 8-oxo-7,8-dihydroadenine (oxo^8A) and 4,6-diamino-5-formamidopyrimidine (FapyA) as the main products [2].

1.1.1.1 5-formyluracil

5-Formyluracil (f^5U) is a common ROS induced base lesion resulting from oxidation of thymine in the 5-methyl group. Although f^5U is formed in DNA at a similar level as guanine oxidized in the 8'position (7,8-dihydro-8-oxoguanine; oxo^8G), the number of studies dedicated to investigate these two lesions are very unequally distributed. Where a great effort has been made to investigate the formation and consequences of oxo^8G, relatively few reports have been published on f^5U. f^5U exists in a keto-enol form of equilibrium as shown in Figure 2, with the different forms exhibiting different base-pairing abilities.

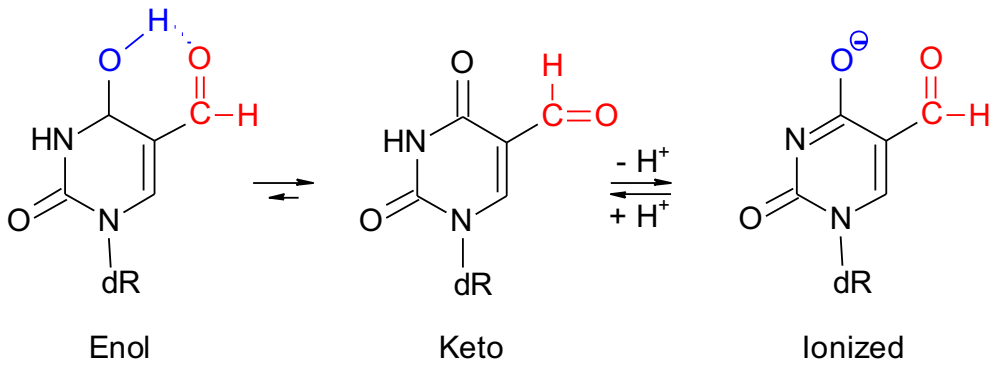


Figure 2: f^5U in enol-, keto- and ionized-form

Adapted from [3]

The keto form ($f^5U^{C=O}$) pairs preferably with adenine (Figure 3a) in a Watson-Crick configuration. However, it can also pair with cytosine (Figure 3c). The ionized form (f^5U^{θ}) pairs preferably with guanine (Figure 3b).

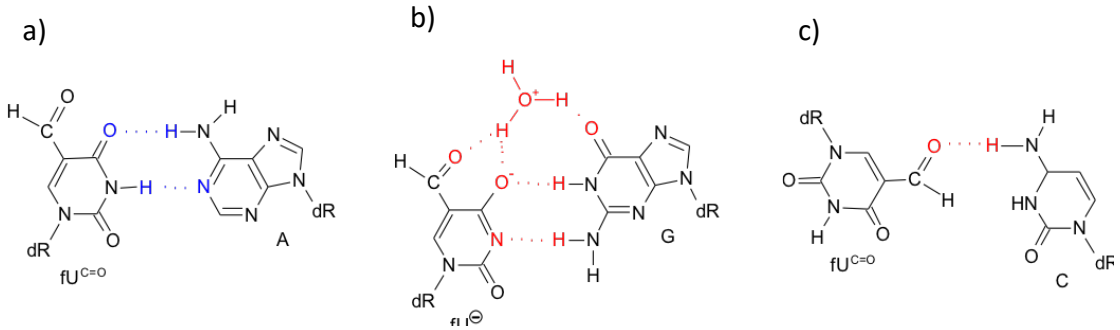


Figure 3: f^5U base pairing abilities

Adapted from [4]

Mutation induction by 5-formyluracil

As a consequence of the diverse base pairing abilities of f^5U it is a mutagenic agent, and a proposed schematic model of the different ways it will induce substitutions in DNA is presented in Figure 4.

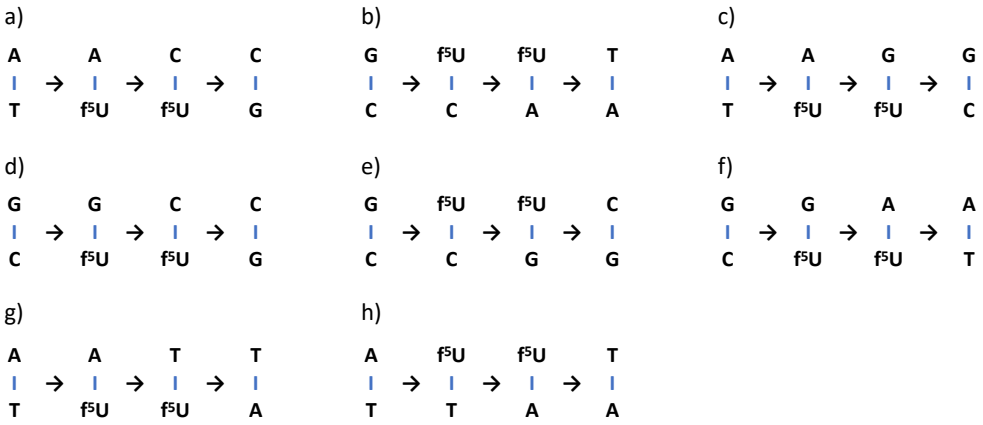


Figure 4: Proposed models for base substitutions induced by f^5U.
a) AT→CG transversion. b) GC→TA transversion. c) AT→GC transition.
d) & e) GC→CG transversion. f) GC→AT transition. g) & h) AT→TA transversion

Adapted from [5]

1.1.2 Methylated bases

Due to the cytotoxicity of methylating agents, they are utilized in cancer treatment, they are also highly mutagenic, where all the bases are susceptible to erroneous methylation by alkylating agents. All the exocyclic oxygens and most of the ring nitrogen atoms can be targeted, in addition to oxygens in the phosphate groups of the sugar-phosphate backbone, however when the ring nitrogen atoms are involved in base pairing in dsDNA, they become quite non-reactive [1, 6]. The alkylating agents are divided in two types by their mode of nucleophilic substitution reactions; S_N1-type agents and S_N2-type agents, where the S_N1-type agents like N-methyl-N-nitrousourea (MNU) alkylate both oxygens and nitrogens, while the S_N2-type agents like methylmethanesulphonate (MMS) mainly alkylates nitrogens [1], as shown in Figure 5.

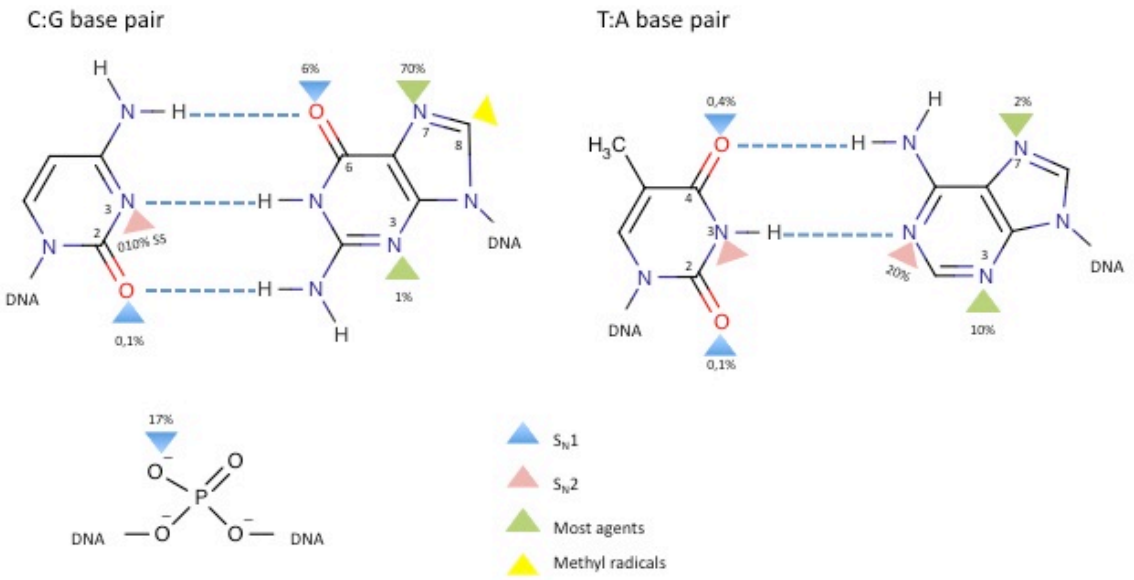


Figure 5: Distribution of methylation sites on the bases and the sugar-phosphate backbone of DNA.

Adapted from [6]

Figure 6 shows the most common methylated bases.

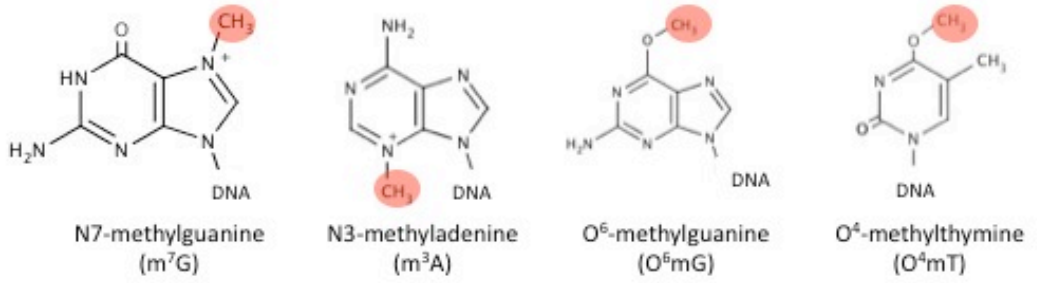


Figure 6: Common methylated bases
Red circle indicates the modified positions

1.1.2.1 Enzymatically methylated bases act as epigenetic regulators of the genome

Enzymatically methylated bases at specific sites in DNA act as epigenetic regulators, where m^5C is so common, constituting up to 4% of the total amount of cytosines in human DNA, that it is often referred to as the "fifth DNA base" [7]. Mostly m^5C acts as a repressor of gene transcription by either blocking transcription factors or by recruiting methyl-binding domain protein that will inhibit gene transcription [1, 8]. The methyl donor for C5-methyltransferase (MTase) is S-adenosylmethionine (SAM), a S_N2-type agent [9]. In addition to its function as a transcriptional repressor in eukaryotes, m^5C also plays an important role in strand recognition during replication, i.e. separating the daughter strand from the parent strand, in prokaryotes. In eukaryotes, the methylation pattern is copied to the unmethylated daughter strand after replication by a maintenance MTase. The methylation pattern of cytosine is not random and mostly occurs in sequences enriched in the CpG dinucleotide (CpG islands), where altered methylation patterns have been implicated in tumorigenesis and developmental issues [10, 11].

While spontaneous hydrolytic deamination of cytosine yields uracil, deamination of m^5C will yield a thymine, where both result in a C→T (G:C→A:T) transition mutation [12]. As described later, the removal of uracil from DNA is very efficient, while the removal of deaminated m^5C is less efficient, probably because the resulting thymine is a native DNA base, making it harder to distinguish from a correctly paired thymine. This could also explain the mutational hotspots seen in CpG sequences. m^5C is also, as all the other bases, a target for attack by ROS, leading to 5-formylcytosine, 5-hydroxymethylcytosine and 5-carboxycytosine lesions.

Bacteria include m^5C as a methylated part of their native DNA, but they also employ N^6-methyladenine ($m^{N6}A$) and *N*4-methylcytosine ($m^{N4}C$). In fact *E. coli* DNA contains twice the amount of $m^{N6}A$ compared to m^5C. While m^5C is thought to have a main role as protection against own restriction enzymes, it is established that $m^{N6}A$ plays a much broader role including involvement in regulation, DNA replication, mismatch repair, strand segregation and regulation of gene expression [13]. Recent studies have also identified $m^{N6}A$ in eukaryotic genomes [14, 15].

1.1.2.2 N^4,5-dimethylcytosine

Because m^5C is common in DNA, it is reasonable to expect it as a target for further methylation (see Figure 5). The Klimasauskas group in Vilnius, Lithuania decided to investigate this further, and in 2002 they published an article where they described successful introduction of N^4,5-dimethylcytosine ($m^{N4,5}C$) *in vitro* [16]. A further investigation to see if the modified base existed *in vivo* was inconclusive, leading to the hypothesis that the $m^{N4,5}C$ residue exists *in vivo*, but is effectively removed by a repair mechanism, this hypothesis was also indirectly supported by the fact that in *E. coli* cells carrying the plasmid used to perform the second N4-methylation, the SOS response was induced [16]. In 2018 our research group in collaboration with the Klimasauskas group published a paper confirming the incision activity of two *E. coli* glycosylases, Nei and Fpg, on short oligonucleotide substrates containing $m^{N4,5}C$ *in vitro* [17].

The Klimasauskas group in Vilnius discovered that they could only methylate cytosine to $m^{N4,5}C$ when it was first methylated in the 5th position of the ring. They explained this by the fact that C5-methylation requires two N4-hydrogen atoms for interactions with the MTase as well as a transient covalent bond in the 6th position (S_N2-type reaction).

The $m^{N4,5}C$ deoxynucleoside has two theoretical conformations, the *cis* conformer which is highly disfavored due to a steric clash between the two methyl groups, and the more likely *trans* conformer, as shown in figure 7. However, the *trans* conformation will not accommodate the normal Watson -Crick base pairing, which could be a reason for efficient *in vivo* removal of $m^{N4,5}C$ from DNA as it is likely to distort its B-helical structure [16].

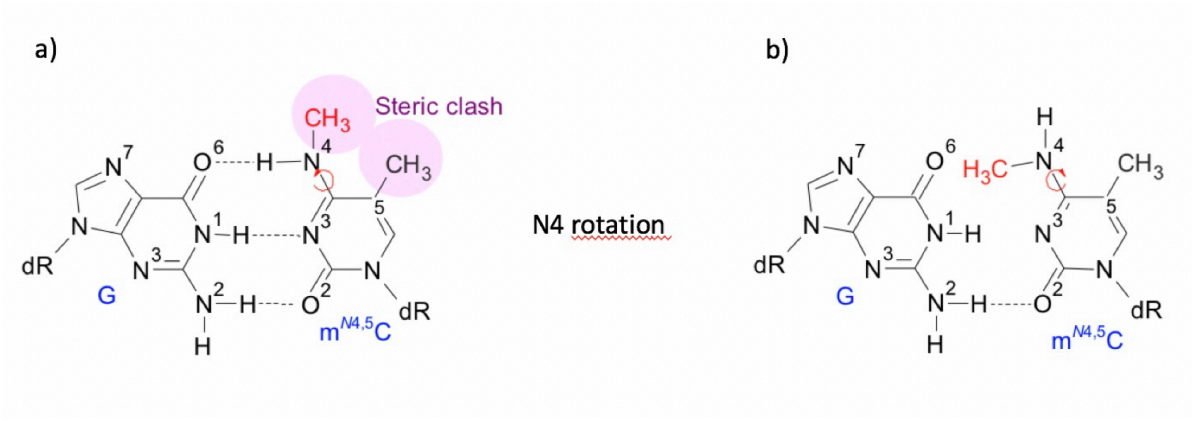


Figure 7:The two isoforms of N^4,5-dimethylcytosine
a: cis-conformer
b: trans-conformer

Adapted from [16]

1.2 DNA Damage repair

It is estimated that there are more than 10.000 oxidative damages occurring per day in human cells, which in addition to hydrolytic damages, alkylation damages and the strand breaks adds up to a massive amount of damages that needs to be handled by the cell [18]. Through time cells have evolved a number of mechanisms to detect and repair the damages that arise from exogenous and endogenous agents, and many of them are highly conserved between all domains of life. In this study the nucleotide excision repair- (NER), and the Base excision repair-pathways (BER) are most relevant and will be explained in some detail. Historically there has been a firm belief that BER handles single base damages, whereas NER tackles the more bulky DNA adducts. However, the NER pathway is also responsible for sorting out single base damages in DNA such as f^5U [19, 20].

1.2.1 SOS Response in *E. coli*

The SOS response is a global defense mechanism against DNA damage in *E. coli*. When the SOS response is induced, the cell cycle will be arrested, and depending on the severity of the damage(s) a number of different repair mechanisms can be initiated.
The SOS pathway in *E. coli* consists of two key regulatory proteins; LexA and RecA, but it is estimated that over 30 genes (SOS genes) are controlled by these regulatory proteins in E. coli, among them genes in the nucleotide excision repair pathway (NER), as well as well as the *lexA*$^+$ and *recA*$^+$ genes themselves.

Mechanism
During normal growth the homo-dimer of LexA acts as a repressor and bind to what is referred to as SOS-boxes of the SOS genes. This causes the transcription of these proteins to stay low. The strength of the repression depends on both the sequence of the SOS-box, and the strength and positioning of the promoter, this feature allows the cell to regulate which response to activate against the damage and at what time. If the cell is subjected to an increased level of DNA damage, the LexA homodimers will dissolve to monomers and no longer repress the transcription of the SOS-genes, in addition the RecA protein will bind to ssDNA and create what is known as RecA-ssDNA filament. This RecA-ssDNA filament promotes cleavage of the LexA homodimers and thus acts as an enhancer of SOS-gene transcription. When the amount of ssDNA in the cell is high, the concentration of RecA-ssDNA is equally high, and the SOS response is induced. The first line of defense are the genes regulated by

weak repression of the LexA homodimers such as the *uvrA*$^{+}$ and *uvrB*$^{+}$ genes, whereas the three polymerases with a less accurate insertion rate Pol III, Pol IV and Pol V have SOS boxes with a stronger binding to LexA homo dimers and can thus be regarded as more of a last resort defense if the former initiated defenses have not been sufficient enough to repair the damages.

1.2.2 Nucleotide Excision Repair

The Nucleotide Excision Repair (NER) mechanism is responsible for recognizing and repairing a wide range of helix distorting lesions to DNA such as carcinogenic cyclobutane pyrimidine dimers (CPD) induced by UV radiation, guanine-cisplatinum adducts formed during chemotherapy, and benzo[α]pyrene-guanine adducts caused by smoking. Helix distorting damages interfere with DNA replication and transcription and is thus an important obstacle to overcome for the cell. It was previously thought that NER did not engage in repairing single-base damages, but it has been proven that this pathway can also repair some oxidation induced damages to DNA, such as oxo^{8}G and Tg, formed by oxidized guanines and thymines respectively.
The NER pathway is simpler in prokaryotes than in eukaryotes, but in short both consists of the following steps:

1. *Recognition* and *verification* of base damage.
2. *Incision* of the lesion containing DNA strand (bimodal or dual incisions).
3. *Excision* of the resulting oligonucleotide fragment.
4. *Repair synthesis,* to fill in the gap from the excised fragment.
5. *DNA Ligation* to seal newly synthesized fragment with original DNA strand.

In this mechanism, after damage recognition and verification, a short segment of the DNA strand is removed both upstream and downstream of the lesion, resulting in a 12-13 nucleotide gap in *E. coli* and a 24-32 nucleotide gap in humans. This is followed by repair synthesis based on the opposite unaffected strand before ligase completes the process.
The NER mechanism has two alternative pathways; Global Genomic NER (GG-NER) and Transcription Coupled NER (TC-NER), where their difference lies in the mode of damage detection. TC-NER is initiated upon stalling of RNA polymerase during transcription, while GG-NER is constantly scanning the genome for damages. Because GG-NER is most relevant to mutation induction and thus this thesis, it will be the one described in detail.

1.2.2.1 Nucleotide excision repair in *E. coli*

In *E. coli* the UvrABC system initiates and carries out the first steps (1 and 2) of the NER mechanism, while the second part (steps 3-5) is carried out by DNA helicase II (UvrD), DNA Polymerase I and DNA ligase. Even though the system consists of three different proteins with different properties, they are as a whole considered as a DNA Damage-Specific Endonuclease. The name “Uvr” reflects the original discovery of them as UV damage repairing proteins (Table 1).

Table 1: Selected properties of the UvrA, B and C proteins.

	Protein		
Property	UvrA	UvrB	UvrC
Mol mass (kDa)	103.9	76.1	66,0
No. of amino acids	940	673	610
DNA Binding	Yes	No	Yes
Nucleotide binding motifs	2	1	0
ATP-ase activity	Yes	No	No
SOS Regulation	Yes	Yes	No
No of molecules/cell*	20 (250)	250 (1000)	10 (10)
(Probable) Function	Molecular matchmaker, initial DNA damage detection, Transcription-coupled repair	Definitive DNA damage recognition, DNA unwinding, 3' incision	Initiation of 3' incision and 5' incision

*Numbers within parentheses reflects values after SOS induction.

For many years there was a common understanding that UvrA forms a homodimer that loosely binds to DNA and scans for damages. Upon contact with such damages $(UvrA)_2$ will recruit UvrB to the site of damage and form a $(UvrA)_2UvrB$ complex more stably bound to DNA. Then $(UvrA)_2$ will disassociate leaving behind a stable UvrB-DNA complex exhibiting high affinity for UvrC, which once recruited to the site of damage binds to UvrB and executes the incisions 8 nucleotides upstream and 4 nucleotides downstream of the damage. Over the years some studies have suggested that the complex formed between UvrA and UvrB is a heterotetramer; $(UvrA)_2(UvrB)_2$. Although no definitive conclusion to this predicament is reached, the mechanism is presented with $(UvrA)_2(UvrB)_2$ as the DNA binding complex in Figure 8 because of increasing evidence for its involvement [21-24]. A more detailed description of each step and the role of the components is as follows.

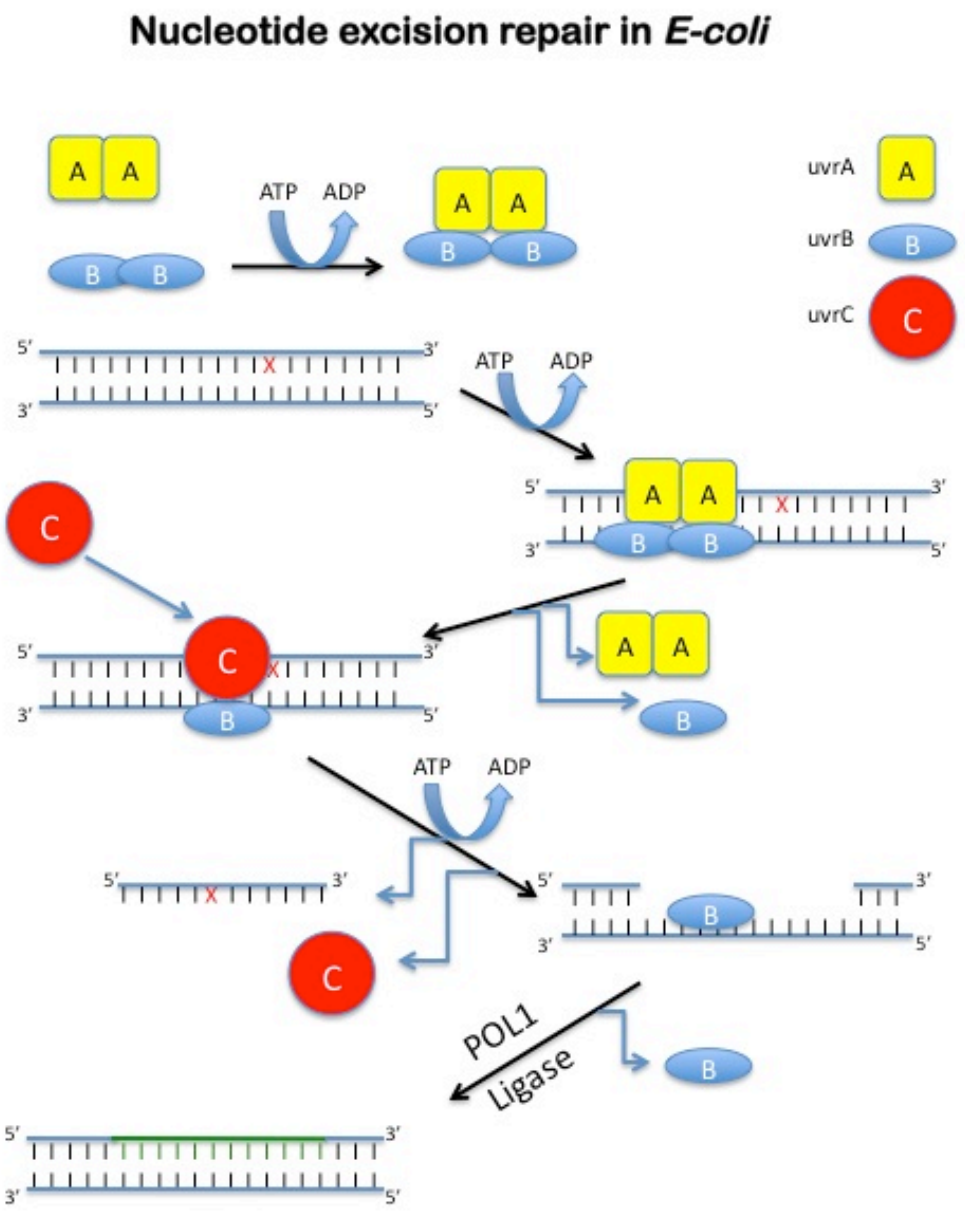


Figure 8: Overview of the NER pathway in *Escherichia coli*

Recognition of base damage

The first player in *E. coli* NER is the UvrA protein encoded by the *uvrA*$^{+}$ gene, which is produced in a ten times higher amount following SOS induction (Table 1). UvrA has several domains and motifs (Figure 9), corresponding to the proposed mechanism for its activity. Among these domains we find two ATP binding domains which are associated with both the need for ATP to execute the conformational change that is required for dimerization [25] and the specificity for binding damaged DNA [26]. The presence of ATP also seem to stimulate the dissociation of UvrA from DNA once UvrB is present [27]. Another important feature of the protein are the two zinc finger domains as well as a consensus helix-turn-helix (HtH) motif. The latter plays an important role in DNA binding, as the C-terminal helix can fit into the major groove of DNA, where it establishes contact between bases of DNA and amino acid residues in the protein. The zinc finger domains are also implicated in the DNA binding [28], and the glycine rich C-terminal end of UvrA is believed to be involved in damage recognition [29].

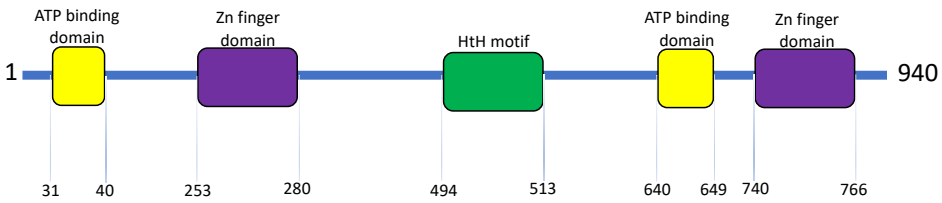


Figure 9: Diagrammatic presentation of the uvrA^{+} gene in *Escherichia coli*

Adapted from [26]

Stable DNA binding and confirmation of damage

Although the UvrA homodimer binds to specific DNA damages, it does not form a stable complex, for stability the mechanism relies upon the second player in the system; UvrB. The UvrB protein is encoded by the *uvrB*$^{+}$ gene which has both SOS-dependent and SOS-independent promoters. The protein is a helicase and has a domain that suggests ATP binding. However, since ATPase activity was not confirmed by the purified protein, it has been suggested that UvrB has a cryptic ATPase, where binding to UvrA and the subsequent domain movement are necessary for its activation [30]. It is worth mentioning that while UvrA can form an unstable and short-lived complex with dsDNA, UvrB will only bind ssDNA in solution in the absence of UvrA [31], which explains why both are needed for damage detection and binding. When scanning DNA for lesions the DNA strand is wrapped around UvrB, and assuming the UvrAB complex is a heterotetramer, the structures of the UvrAB-DNA complex ensures damage detection in both strands by alternate scanning of the two strands [22]. This is facilitated by ATP hydrolysis in the first subunit that will release the DNA, followed by ATP binding in the second unit that will ensure the DNA wrapping around the latter unit. A proposed model for this damage detection by UvrB involves the protein trying to flip individual nucleotides out of the helix, when no damage is present nucleotides are held in place by stacking forces, and this prevents the necessary unwinding of the helix by the ß-hairpin structure of UvrB [32]. If a damage is detected and verified by UvrB, the UvrB-DNA pre-incision complex is formed by insertion of the UvrB ß-hairpin followed by the dissociation of the UvrA dimer. Whether the UvrB-DNA pre-incision complex consists of one or two UvrB proteins is not determined at this point, one theory proposes that upon detection and verification of a damage the UvrB protein on the opposite strand is obsolete, and will disassociate before the third player in the mechanism; UvrC, is bound to the pre-incision complex. This theory is somewhat supported by the fact that the C-terminal domain of UvrB is required for dimerization [33], but also for binding to a homologous amino acid sequence within UvrC [34].

Incision

In prokaryotic NER the incision is bimodal and executed by the third player in the mechanism; the UvrC protein. This protein is encoded by the *uvrC*$^{+}$ gene which unlike the two other Uvr proteins is not linked to the SOS response. Mapping of the *uvrC*$^{+}$ gene reveals two catalytic domains separated by a binding domain for UvrB, as well as two helix-hairpin-helix (HhH) motifs at the C-terminal end of the protein as illustrated in Figure 10.

The first incision is the cleavage of the phosphodiester bond 4 or 5 nucleotides 3' to the damage [35], this step requires the interaction between the C-terminal domain of UvrB and the homologous UvrB binding domain of UvrC [36]. The 3' incision requires the presence but not hydrolysis of ATP, apparently because binding of ATP is required to execute a conformational change in the complex to facilitate the incision [35]. The second incision is made by cleavage of the 8th phosphodiester bond 5' to the damage, this incision does not require the interaction between domains of UvrB and UvrC, nor the presence of ATP [26]. After both incisions are carried out by UvrC the fragment is excised by DNA helicase II (sometimes named UvrD).

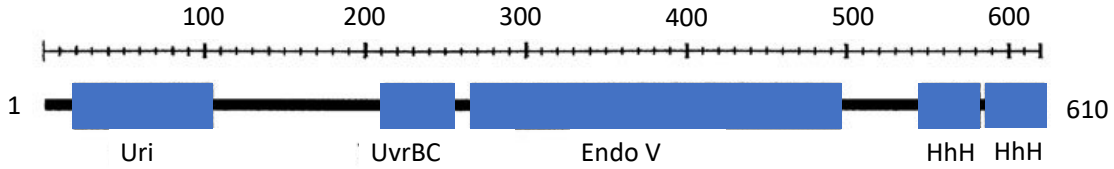


Figure 10: Diagrammatic presentation of the uvrC^{+} gene in *Escherichia coli*.
N-terminal catalytic Uri domain involved in 3' incision, UvrBC domain involved in binding with UvrB, catalytic Endo V domain involved in 5' incision and the two C-terminal HhH motifs also involved in 5' incision.

Adapted from [37]

Excision

DNA helicase II (UvrD) is encoded by the *uvrD*$^{+}$ gene and is induced by the SOS response [38, 39]. In the context of NER it is responsible for excising the oligonucleotide fragment containing the damage, as well as releasing UvrC from the incision complex. As no other DNA helicases in *E.coli* can act as a substitute in this process, it is presumed that this is a highly specific function of DNA helicase II [40]. It remains unknown whether DNA helicase II interacts with the post-incision complex itself, or the generated nicks. However, it is clear that UvrB is not released at this point in the process, but will remain bound to the gapped DNA until repair synthesis has taken place [40].

Repair synthesis of DNA

After the damaged fragment of DNA is removed by UvrD, DNA Polymerase I (Pol I) starts synthesis from the 3' OH terminus generated at the 5' incision site. As Pol I progress it will displace the bound UvrB protein, eventually releasing it from DNA. Some studies have challenged the view that Pol I is the only DNA polymerase that can engage in repair synthesis following the initial steps of NER. However, Pol I was concluded as the preferred polymerase [41, 42].

DNA ligation

After Pol I has completed DNA synthesis, DNA ligase will seal the nick 3' to the newly synthesized strand. *E. coli* contains only one DNA ligase encoded by the *ligA*$^{+}$ gene, which depends on NAD^{+} as energy source for the end-joining process

1.2.2.2 Nucleotide excision repair in mammalian cells

Nucleotide excision repair in mammalian cells follows the same basic steps as in *E. coli*, but the complexity of the mechanism and number of involved players are much higher. For further details regarding NER in mammalian cells the following references are recommended: [43, 44]
However, the severe consequences of dysfunctional NER in humans is worth mentioning, as it highlights the importance of this mechanism and the need for understanding it, at a more comprehensible level to the reader. Three hereditary diseases are identified to be caused by defects in the NER pathway in humans; Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. To various degrees these conditions result in high susceptibility to skin cancer, neurological disorders and premature ageing.

1.2.3 Base Excision Repair

The existence of base excision repair (BER) was discovered in 1974 by Tomas Lindahl when he searched for an enzyme that would remove deaminated cytosine (uracil) from the genome, what he then discovered was however not a nuclease as he expected, but an enzyme that cleaved the bond between the damaged base and the deoxyribose in the sugar-phosphate backbone of DNA. Thus, *E. coli* uracil-DNA glycosylase (Ung) was the first DNA glycosylase discovered and laid the ground for defining the BER pathway in molecular detail in both prokaryotes and eukaryotes [45]. We now know that this pathway not only relieve genomic uracil from DNA but also includes repair of most non-helix distorting single base damages in DNA caused by oxidation, alkylation and deamination [46]. The basic steps of this mechanism are essentially similar as for NER, with damage recognition as the first step, followed by excision, incision, repair synthesis and ligation. As indicated for uracil, the first step is carried out by a class of DNA repair enzymes called DNA glycosylases, which leave behind an AP site. While NER excises a small fragment of DNA, BER will remove the damaged base only (short patch repair) or a very short segment of DNA (long patch repair, usually a few bases).

1.2.3.1 DNA Glycosylases

The BER pathway is initiated by a DNA glycosylase, which is defined as either monofunctional or bifunctional, depending on its catalytic mechanism. The monofunctional glycosylases cleave the N-glycosidic bond between the damaged base and the deoxyribose moiety [47, 48], leaving behind an AP site. The bifunctional glycosylases exhibit AP-lyase (β-elimination) activity in addition to the glycosylase activity, which yields a 3' α,β-unsaturated aldehyde adjacent to a 5' phosphate (Figure 11). This contrasts with the AP enconuclease following a monofunctional glycosylase, which leaves behind a free 3’-OH end ready for repair replication and a 5’ deoxyribosephosphate remnant. Some bifunctional glycosylases can further perform a ð-elimination converting the 3’ aldehyde to a 3’ phosphate, but in both cases further processing by either a 3’-phosphodiesterase (3’ α,β-unsaturated aldehyde) or a 3’-polynucleotide phosphatase (3’ phosphate) is required to obtain a free 3’-OH end, in order for the repair synthesis to take place. Presently, at least 8 *E. coli* and 13 human DNA glycosylases are identified (Table 2).

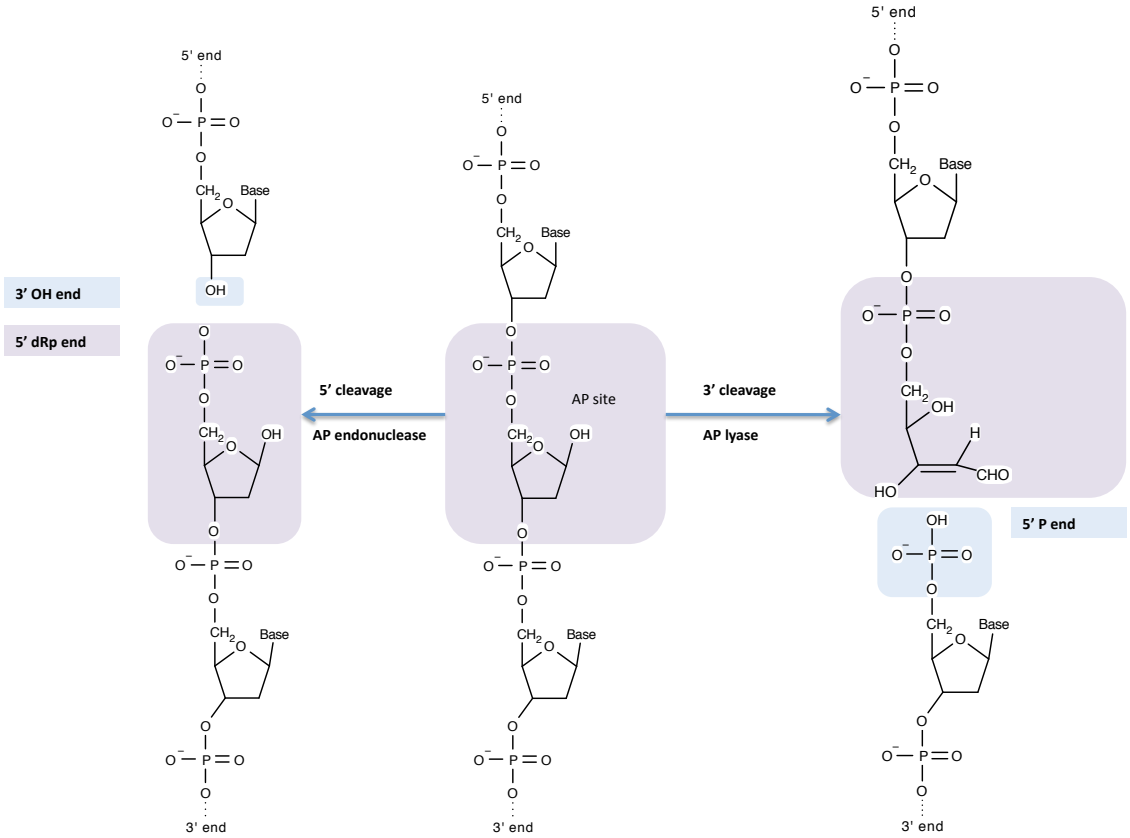


Figure 11: Detailed description of the AP site and cleavage by AP endonuclease and AP lyase activities. The middle represents a single strand of DNA with an AP site. To the left: cleavage on the 5' side of the AP site by an endonuclease resulting in a 3' OH terminus and a 5'dRp residue. To the right: Cleavage on the 3' side of the AP site by an AP lyase results in a 5' phosphorylated end and a 3' end with a 3' unsaturated aldehyde.

Adapted from [49]

Table 2: Identified *Human* and *E. coli* glycosylases and their known substrates

Base damage	Name	Gene	Mono/bifunctional	Known substrates
			Human	
Deaminated	hUNG	*UNG*	M	ssU, U:G, U:A, h^5U, isodialuric acid, alloxan, 5-fluoroU
	hSMUG 1	*SMUG 1*	M*	ssU, U:A, U:G, f^5U, hm^5U, h^5U, εC:G, 5-fluoroU:A
	hUDG2	*UDG2*	m	U:A
	hMBD4 (MED1)	*MBD4*	M	U:G, T:G, hm^5U in CpG,f^5U, m^5C, εC in ss and ds DNA, 5-fluoroU
	ThDG	*TDG*	M	U:G, T:G, T:C, T:T, f^5U:G, f^5U:A, hm^5U:G, hm^5U:A, εC in ss and ds DNA, f^5C, ca^5C, 5-fluoroU
Oxidized	hOGG 1	*OGG 1*	M/B, ß-elimination	fapyG:C, oxo^8G, oxo^8A:C, mfapyG:C
	hNEIL 1	*NEIL*	B, ß/∂-elimination	Tg, fapyG, fapyA, mFapyG, oxo^8G, h^5C, h^5U, dHU, dHT, Sp and Gh in ss and ds DNA
	hNEIL 2	*NEIL*	B, ß/∂-elimination	h^5U, dHU, h^5C, dHT, Tg, dHU, oxo^8G,
	hNEIL 3	*NEIL*	M/B, ß/∂-elimination	FapyG, FapyA, Sp and Gh in ss DNA,
	hNTH 1 (NTHL1)	*NTHL 1*	B, ß-elimination	
Alkylated	hMUTY	*MYH*	M	A:G, oxo^8G, A:C mismatches
	hMPG	*MPG*	M	m^3A, m^3G, m^7G, hypoxanthine, εA, EA, $m^{N6}A$, m^7A, oxo^8G
			Escherichia coli	
Oxidized	Fpg	fpg^+	B, ß/∂-elimination, dRP lyase	Oxo^8G, Oxo^8A, FapyG:C, FapyG:A, mFapy, FapyA, h^5C, h^5U, Tg, dHT, hmh, hm^5U, f^5U, Ug, Gh, AP sites, εA, ring opened εA
	Nei	nei^+	B, ß/∂-elimination, dRP lyase	Oxo^8G, FapyA, FapyG, Tg, dHT, 5-OH-dHU, hmh, f^5U, hm^5U Ug, h^5U, h^5C, dHC, dHT, dHU, Gh, Sp, AP sites, Th^5, urea, dhC, dhU, Uh^5, oxanine, xanthine, β-ureidoisobutaric acid
	Nth	nth^+	B, ß-elimination	FapyG, Gh, Sp, FapyA, AP sites, Tg, 5-OH-dHT, 6-oh-dHT, dHT, Th^5, hmh, hm^5U, 5-OH-dHU, 6-OH-dHU, h^5C, dhC, 6-OH-dHC, Ug, h^5U, dHU, dhU, mFapy, Cg, f^5U, urea,ring opened εA
	MutY	$mutY^+$	M	A:G, A:C, A:oxo^8G
Methylated	Tag	tag^+	M	m^3A, m^3G, m^7G
	AlkA	$alkA^+$	M	m^3A, m^3G, m^7A, m^7G, hm^5U, f^5U, $m^{O2}C$, $m^{O2}G$, EA, εA
Deaminated	Mug	mug^+	M	dsU, U:G, T:G, ssT, hm^5U:A, hm^5U:G, f^5U:A, f^5U:G, h^5C:G, h^5U:G
	Ung	ung^+	M	ssU, U:G, U:A, h^5U, dhU

Parentheses denotes most common synonyms.
* Although hSMUG1 is historically viewed as a monofunctional glycosylase, a paper was published in 2019, stating that it also has incision activity and further processing activity [50].

Adapted from [26, 51]

In this thesis, three glycosylases were investigated for activity against $m^{N4,5}C$ lesions; *E. coli* MutY, hMPG and hSMUG (25-270). A more detailed overview of the investigated glycosylases will follow.

MutY

The most important task of the *E. coli* Adenine DNA glycosylase (MPG) is to catalyze the excision of A from oxo^8G:A mispairs. This acts as a backup mechanism in the cases where oxo^8G (usually paired with C) is not removed prior to replication by Fpg. Since replicative polymerases sometimes inserts an A opposite oxo^8G instead of a C. MutY will initiate BER in the undamaged strand, yielding an oxo^8G:C pair which is a recognizable and a suitable substrate for Fpg to initiate BER in the damaged strand.

hMPG

N-Methylpurine-DNA glycosylase (MPG) has similar functions to *E. coli* AlkA and utilizes the same base flipping principal but the two are not related. MPG has a broad substrate specificity and efficiently excises both neutral substrates as well as positively charged methylated bases from DNA [26]. MPG show very low activity against unmodified purines, by promoting acid-catalyzed excision of purines from DNA while retaining a selective filter against the 6-amino and 2-amino groups of purines, broad substrate specificity is obtained without unwanted cleavage of undamaged DNA [26].

hSMUG1

The constitutive enzyme human single-strand-selective mono-functional UDG (hSMUG1) is a member of family 3 of UDGs. Unlike hUNG which is upregulated during replication, hSMUG1 is continuously expressed in non-replicating cells [52, 53].
Historically hSMUG1 has been classified as a monofunctional glycosylase, yet a recent study has confirmed strand incision and processing activity in this enzyme, yielding a 3'-α,β-unsaturated aldehyde and a 5'-phosphate, and also further processing of the 3'-α,β-unsaturated aldehyde to 3'-phosphate [50]. The known substrates of hSMUG1 are listed in Table 2. The truncated version used in this thesis has the same activity but lacks the first 24 amino acids.

1.2.3.2 Base excision repair in *E. coli*

Damage recognition and excision of altered bases

As mentioned above, the first glycosylase identified was the *E. coli* UDG Ung. Over the years several other uracil-DNA glycosylases have been identified and classified into six families where Ung is a member of family 1. The active site pocket in this family is highly conserved between the higher organisms and very specific to accommodate uracil binding, and disallow binding of most other bases [54]. Another UDG in *E. coli* is Mug, which is a member of family 2. In addition to being a mismatch-specific glycosylase for guanine (removing uracil and thymine opposite guanine), this enzyme has specificity for removing lesions involving f^5U, 5-hydroxycytosine (h^5C) and 5-hydroxymethyluracil (hm^5U) from DNA.

In *E. coli*, alkylated bases are removed by 3-methyladenine glycosylase I (Tag) and 3-methyladenine glycosylase II (AlkA). AlkA has a wide substrate specificity and is part of the adaptive response to alkylation, but is only responsible for removing about 10% of the erroneously methylated bases unless induced by alkylating agents, after induction the percentage rises to 50-70% [48]. Tag on the other hand is a constitutive enzyme which has a much more specific substrate recognition and is responsible for removing 3-methyladenine (m^3A), and 3-methylguanine (m^3G) [55], and has a strong preference for dsDNA.

Oxidized bases in DNA are in *E. coli* removed by endonuclease III (Nth), formamidopyrimidine DNA-glycosylase (Fpg) and endonuclease VIII (Nei), which are all bifunctional glycosylases. It is interesting to note that although Nei has significant sequence homology to Fpg, its substrate specificity has

more overlap with Nth. MutY is also responsible for removing some oxidized bases as mentioned above.

Incision and end-processing

AP sites generated by monofunctional glycosylase excision are cytotoxic by inhibiting DNA replication and transcription, mutagenic and fragile in terms of strand break generation. DNA ends generated after excision and incision by the bifunctional glycosylases are either not yet suitable for repair synthesis (no free 3'-OH to act as a primer for Pol I) or ligation (a 5'-dRP instead of a 5'-P). Incision of the DNA strand is carried out by an AP endonuclease (APE), in *E. coli* either Exonuclease III (Xth) or Endonuclease IV (Nfo). They are both 5' AP endonucleases and very diverse enzymes that also has 3'-phosphodiesterase and 3'-phosphatase activity that can remove 3' unsaturated aldehyde and 3' phosphate, generated through ß-elimination and ð-elimination by the bifunctional glycosylases Nth and Nei/Fpg, respectively [56]. Xth and Nfo have significant overlap in the substrate specificity, and a main feature that separates them is the 3'→5' exonuclease activity of Xth. Xth is responsible for 90% of the total removal of 3'-blocking ends in *E. coli* DNA [57], however, it is suggested that Nfo may recognize some lesions overlooked by Xth [58]. Although the bifunctional glycosylases exhibit AP lyase activity, they cannot completely substitute the functions of the APEs, and in many cases a bifunctional glycosylase will leave behind an AP site to be further processed by an APE. Both Xth and Nfo use base flipping and hydrolytic attack facilitated by metal ions to perform the incision 5' to the AP site, but where Xth requires a Mg^{2+} ion, Nfo has a tightly bound trinuclear cluster of Zn^{2+} ions in its active site.

Repair synthesis and ligation

Pol I is identified as the most important gap filling polymerase associated with the BER pathway [59], where usually only the single altered base is replaced (short patch repair). However, the 5'→3' exonuclease together with the polymerase activity of Pol I can sometimes digest and replace the neighboring bases for a short stretch downstream of the damaged base (long patch repair, usually 2-8 bases). The enzymes involved in removing the 5'-dRP seems to determine whether the BER pathway continues via short patch or long patch repair. In short; if the 5'-dRP is removed by Fpg, Nei or RecJ or by the 5'→3' exonuclease activity of Pol I, short patch BER will follow [1], if the 5'dRP is not removed prior to initiation of repair synthesis, long patch repair will follow, where Pol I will displace the dRP containing strand via strand displacement [59]. The ligation process for BER in *E. coli* is the same as for NER.

An overview of the entire BER pathway in *E. coli* is presented in Figure 12.

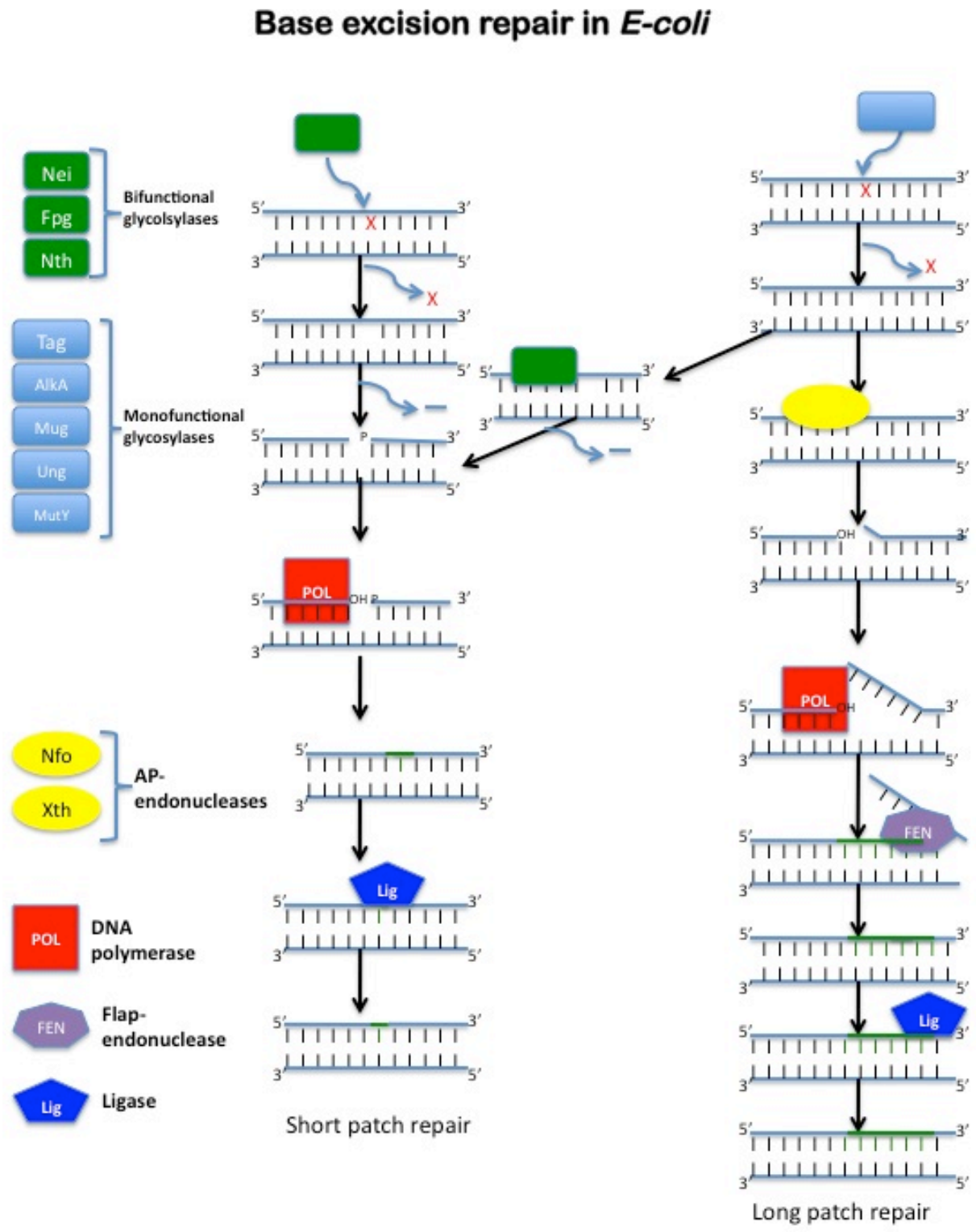


Figure 12: Overview of the BER pathway in *Escherichia coli*.

1.2.3.3 BER in mammalian cells

As with NER, the mechanism of BER in mammalian cells is comparable to prokaryotic BER, although more intricate and with a higher number of players involved. A complete overview of the mechanism is presented in Figure 13.

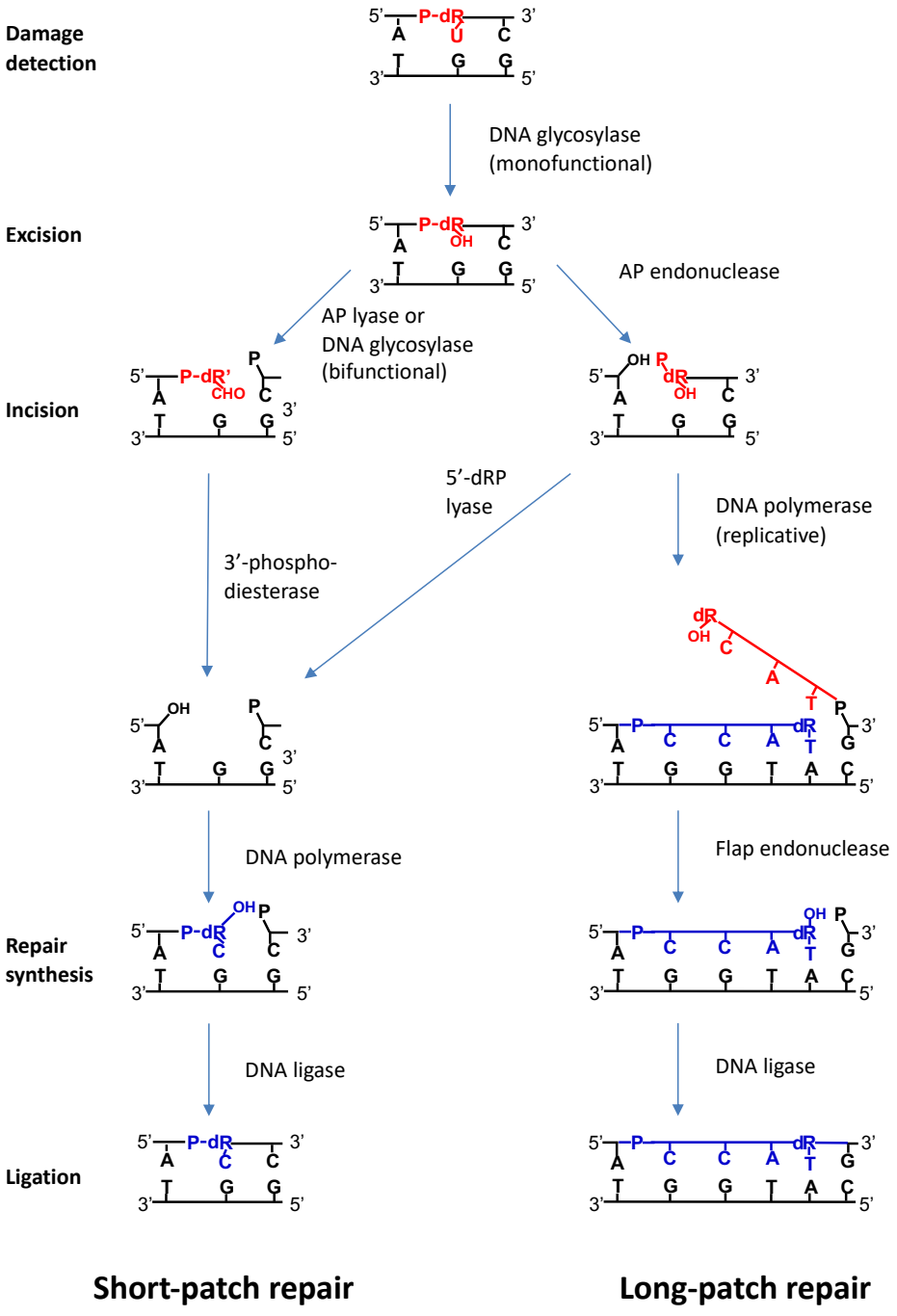


Figure 13: Overview of the BER pathway in mammalian cells

Adapted from [60]

1.3 Aim of study

The original aim of this study was to finish identifying *E. coli* and human DNA glycosylases with the ability to initiate repair of $m^{N4,5}C$ in DNA *in vitro*, and then extend this work to $m^{3,5}C$. However, our collaborators producing DNA with $m^{3,5}C$ experienced unforeseen problems and could not deliver the appropriate substrate. Consequently, the focus of the thesis was shifted half way through the work.

Since I had previous experience (bachelor thesis) with investigating the effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine (f^5dU) in *E. coli*, a larger project not yet finished, we decided to finish the latter project as a part of this master thesis work.

It has been a common belief that the BER pathway is the primary repair mechanism for DNA base lesions induced by oxidation including f^5U, but there is now *in vitro* evidence that the NER pathway also plays a role in the repair of f^5U lesions in mammalian cells [19]. Another study within our group, supplementing *uvrA*$^-$ cells with 5-formyldeoxyuridine during growth followed by analyzing the mutations formed indicated that the same is true for *E. coli*. To determine whether it is the UvrA protein alone that is somehow involved in repairing or recognizing f^5U lesions, or if it is the entire UvrABC system or the NER pathway, this study was started on uvrB$^-$ cells and uvrC$^-$ cells in 2014/2015 and performed using the exact same protocol as with the *uvrA*$^-$ cell study. Together these three studies are closely linked regarding the possible understanding of the role of UvrABC/NER in f^5dU-mediated mutagenesis in *E. coli*.
In all these studies f^5dU was added to the growth medium of exponentially growing bacteria. Like other thymidine analogues f^5dU is thought to be actively transported in to the *E. coli* cells, and subsequently converted by *in vivo* enzymes into 5-formyl-2'-deoxyuridine triphosphate (f^5dUTP) which is then used as a substrate by Pol I.

2 Materials and methods

Complete list of all materials and detailed protocols can be found in appendix A1 for the section regarding effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine in *E. coli*, appendix A2 for the section regarding DNA Glycosylase activities for N^4,5-dimethylcytosine and appendix A3 for the section regarding production and purification of hSMUG (25-270).

2.1 Effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine in *Escherichia coli.*

2.1.1 Mutagenesis

Materials

f^5dU was prepared as described (Ono, Okamoto et al., 1994).
Rifampicin was obtained from G Biosciences (Cat.#: RC-193, Lot#: 152312) and Alfa Aesar (Cat#: 60836, Lot#: T31E022)

Bacterial cells

Two different strains of *Escherichia coli* K-12 were used in this part of the thesis; AB1884 and AB 1885 (Both obtained from Coli Generic Stock Center, University of Yale). Freeze cultures were stored at -80°C.

Buffers

Buffers and media are listed in Table 3.

Table 3: Buffer and media used in mutagenesis assay

Buffers/solutions	Composition
10× Buffer A	600 mM K_2HPO_4, 330 mM KH_2PO_4, 75 mM $(NH_4)_2SO_4$, 17 mM $C_6H_5Na_3O_7$ x $2H_2O$
A media	60 mM K_2HPO_4, 33 mM KH_2PO_4, 7.5 mM $(NH_4)_2SO_4$, 1.7 mM $C_6H_5Na_3O_7$ x $2H_2O$, 0.1 M $MgSO_4$, 0.25% glucose, 8 mg L-aminoacids, 1 mg vit B_1

Mutagenesis was performed essentially as described [61], where all growth was carried out at 37°C. Cells were taken from freeze culture to LB media (5 ml) and grown overnight, then spread out on a LB agar plate for further growth followed by short-term storage at 4°C. To keep a fresh stock of bacterial colonies, one colony was once a week transferred and spread out to a new LB plate and grown overnight before short term storage at 4°C.
For each assay 6 colonies were chosen from the LB stock plate and grown overnight in 2 ml A medium [1 × Buffer A containing 1 mM $MgSO_4$, 0.2% (w/v) glucose, 0.04 mg/ml L-amino acids (Thr, Arg, Pro, Leu, His) and vitamin B_1 (5 µg/ml)], before cell number was determined by OD at 600 nm (one control with no colony was also prepared and used as blanking media for the OD measurements).
The overnight culture (10 µl) was diluted in 1×A buffer to 200 000 cells/ml, before the diluted culture (293 µl) was added to A medium (6.5 ml) and grown for 2 h in falcon tubes to adapt cells and increase cell number. Each culture was then divided to three culture tubes (2 ml each), before addition of f^5dU (0.1 mM) (control samples were grown without f^5dU, to determine spontaneous mutagenesis for each strain), this point defined the start of the mutagenesis experiment/culture, and the cultures were grown for 45-48 hours before being plated for Rif^R mutants on minimal agar plates containing rifampicin (150 µg/ml) and viable cells.

2.1.2 DNA- extraction, -amplification and -sequencing

Materials

Forward primer: 5'-GCCAAGCCGATTTCC-3' (F-1021) (DNA Technology A/S: 381044)
Reverse primer: 5'-GTATTCGTTAGTCTG-3' (R-1021) (DNA Technology A/S: 381045)

Buffers

Buffers and solutions are listed in Table 4.

Table 4: Buffers used for DNA- extraction, -amplification and -sequencing

Buffer	Composition/Provider
5 × colorless GoTaq® Flexi buffer	Promega Ref#: M890A, Lot#: 0000129120
5 × Green GoTaq® Flexi buffer	Promega Ref#: M891A, Lot#: 0000135046
1 × TAE running buffer	40 mM TRIS, 1 mM EDTA pH 8.0, 0.12% acetic acid

Mutant colonies were transferred to LB media (2 ml) containing 150 µg/ml rifampicin (only one mutant from each start culture was chosen) and allowed to grow for 5-7 days before chromosomal DNA was extracted for polymerase chain reaction (PCR), by heating 5 µl of culture in 100 µl of sterile water at 100°C for 5 min, followed by cooling on ice, centrifugation and collection of the supernatant.
The rpoB Rif^R region was amplified by PCR using the forward primer 5'-GCCAAGCCGATTTCC-3' (F-1021) and the reverse primer 5'-GTATTCGTTAGTCTG-3' (R-1021) (0.2 pmol/µl each) using GoTaq® HotStart Polymerase (Promega) as recommended by the manufacturer (final volume 50 µl).
To verify presence and size, the PCR products were run on a 1% agarose gel for 30 min at 100 V, before determining concentration and quality with nanodrop measurement (Thermo Scientific NanoDrop 1), and when necessary, purified using NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel).
DNA sequencing (using F-1021 as primer) was performed by GATC Biotech, Cologne, Germany (using Applied Biosystems 3730xl DNA Analyzer).

2.2 DNA Glycosylase activities for N^4,5-dimethylcytosine

Materials

All substrates used are listed in Table 5.

Table 5: Examined substrates and control substrates used for glycosylase activity testing

	Substrate DNA	Oligo	Sequence 5'-3'
Examined substrates	$m^{N4,5}C$:C	Fw: (Cy3) $m^{N4,5}C$ Rev: C	Fw: C*G*G*TGAAGTAC[$m^{N4,5}C$]AGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTCGTACTTCA*C*C*G
	$m^{N4,5}C$:A	Fw: (Cy3) $m^{N4,5}C$ Rev: A	Fw: C*G*G*TGAAGTAC[$m^{N4,5}C$]AGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTAGTACTTCA*C*C*G
	$m^{N4,5}C$:G	Fw: (Cy3)$m^{N4,5}C$ Rev: G	Fw: C*G*G*TGAAGTAC[$m^{N4,5}C$]AGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTGGTACTTCA*C*C*G
	$m^{N4,5}C$:T	Fw: (Cy3) $m^{N4,5}C$ Rev: T	Fw: C*G*G*TGAAGTAC[$m^{N4,5}C$]AGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTTGTACTTCA*C*C*G
Control substrates	A:G	Fw: A Rev: G	Fw: C*G*G*TGAAGTACAAGGAAGCGATTTCGA*C*C*C Rev:G*G*G*TCGAAATCGCTTCCTGGTACTTCA*C*C*G
	εA:T	Fw: (Cy3) EA Rev: T	Fw: C*G*G*TGAAGTAC[iEth-dA]AGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTTGTACTTCA*C*C*G
	U:G	Fw: (Cy3)U-30 Rev: G	Fw: C*G*G*TGAAGTACUAGGAAGCGATTTCGA*C*C*C Rev: G*G*G*TCGAAATCGCTTCCTGGTACTTCA*C*C*G

Enzymes

Investigated enzymes are listed in Table 6.

Table 6: Enzymes investigated for excision activity against $m^{N4,5}C$

Enzyme	Supplier/Catalog/Lot no/	Dissolved in	Reaction buffer	Control substrate
MutY	Trevigen/4000-500-EB/42623E18	20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT, 50% (v/v) glycerol	1 x REC™	A:G
hMPG	New England BioLabs Inc/M0313S/0021707	10 mM Tris-HCl pH 7.5, 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 50% glycerol, 0.5% Tween® 20, 0.5% IGEPAL® CA-630	1 x ThermoPol® buffer	EA:T
hSMUG (25-270)	Expressed and purified in lab	50% Equilibration buffer (50 mM TRIS pH 7.5, 300 mM NaCl) and 50% glycerol	5 x HEPES with 5 mM DTT	U:G

Buffers
The buffers used in this assay are listed in Table 7.

Table 7: Buffers used for glycosylase activity assay

Buffer	Composition
5 × HEPES	225 mM 4-(2-hydroxyrthyl)-1-piperazineethanesulfonic acid pH 7.5, 10% glycerol, 2 mM EDTA
10 × REC	100 mM HEPES-KOH pH 7.4, 1 M KCl, 100 mM EDTA 1.0 mg/ml BSA
10 × ThermoPol® buffer	200 mM Tris-HCl pH 8.8, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$, 1% Triton®-X-100,
1 × TBE running buffer	89 mM Tris base, 89 mM boric acid, 2 mM EDTA pH 8
formamide gel loading buffer	80% formamide, 1 mM EDTA and 1% (w/v) blue dextran

Substrate oligonucleotides containing $m^{N4,5}C$ opposite A, C, G, or T (Table 5) were incubated with DNA glycosylases at 37°C for 1 h in appropriate reaction buffers according to Table 7. Positive control of enzyme activity was verified by control substrates listed in table 5 incubated under the same conditions. Reactions were terminated by the addition of Stop solution (20 mM EDTA, 0.5% (w/v) and sodium dodecyl sulfate (SDS)) and proteinase K (10 µg), and further incubated for 10 min at 37°C. t-RNA (16 µg) was added as a carrier for nucleic acid precipitation and the DNA precipitation was carried out in 96% ethanol containing 0.1 M sodium acetate (CH_3COONa) at -20°C overnight. Centrifugation (13 000 rpm, 4°C, 30 min) was used to collect the DNA pellets before they were washed with 70% ethanol (-20°C). Ethanol disrupts sample loading into the wells of the SDS-PAGE gel, therefore pellets were dried on ice in ventilation cabinet (20 min) prior to resuspension in 0.1 M NaOH (10 µl) and heat treatment (90°C, 10 min). All tested enzymes were monofunctional and activity was confirmed by NAOH and heat mediated incision following the creation of AP sites [53], for negative controls enzyme specific reaction buffer was added in the same amount as enzyme to obtain equal sample volumes. Samples were then mixed with formamide gel loading buffer (10 µl), heated for 5 min at 90°C and immediately cooled on ice. Samples (5 µl) were loaded to prewashed wells and run on freshly made denaturing PAGE gels (20% (w/v) polyacrylamide gel containing 8 M urea for 2 h at 200 V in RT and darkness using 1 × TBE as running buffer.

2.3 Production and purification of hSMUG (25-270)

2.3.1 Making competent cells

Bacterial cells
E.coli BL21 (DE3) (from in-house stock), freeze cultures stored at -80°C

A single colony of *E. coli* BL21 (DE3) was inoculated in LB media (3 ml) and grown overnight at 37°C with vigorous shaking. The overnight culture was then transferred (200 µl) to LB media (25 ml), and grown at 37°C with vigorous shaking, until the culture reached an OD_{600} of 0.3 – 0.5. Once the culture had reached the desired value for the OD_{600} measurement, the culture was placed on ice for 10 min. The culture was then divided to round bottom falcon tubes (4 tubes, 6 ml each), before they were centrifuged for 10 min at 4000 x g and 4°C. Supernatant was decanted and discarded before the pellets were resuspended in 100 mM $CaCl_2$ (3 ml, 4°C) and left to incubate on ice for 30 min.
Cells were harvested by centrifugation (10 min, 4000 x g and 4°C), supernatant was again decanted and discarded and the pellet was resuspended in $CaCl_2$ (400 µl). Cells not immediately used were snap frozen in liquid nitrogen, and stored for later use at -80°C.

2.3.2 Transforming bacteria

Plasmid

pETM-11 hSMUG (25-270). (a kind gift from prof. Hilde Nilsen)

Plasmid (50 ng) was added to competent cells (200 µl) before incubation on ice for 30 min. The tubes were then placed in a 42°C water bath for exactly 90 seconds and then immediately placed on ice to cool down. LB media (1 ml) was added to the tubes before 1 h incubation at 37°C with vigourous shaking followed. The transformation solution was then plated (200 µl) on a LB plate containing Kanamycin (50 µg/ml) and grown overnight at 37°C.

2.3.3 Autoinduction

A single colony from the previous step was inoculated in ZYM-5052 (500 ml) containing kanamycin (50 µg/ml), and then incubated at 28°C for 24 hours with vigorous shaking (220 rpm). Cells were harvested through centrifugation at 6000 rpm for 20 minutes at RT.

2.3.4 Affinity purification

Buffers

Buffers are listed in table 8.

Table 8: Buffers used for production and purification of hSMUG (25-270)

Buffer	Composition
Lysis buffer	50 mM TRIS pH 7.5, 300 mM NaCl, 5% glycerol
Equilibration buffer	50 mM TRIS pH 7.5, 300 mM NaCl
Wash buffer	50 mM TRIS pH 7.5, 300 mM NaCl, 10 mM imidazole
Elution buffer 1	50 mM TRIS pH 7.5, 300 mM NaCl, 100 mM imidazole
Elution buffer 2	50 mM TRIS pH 7.5, 300 mM NaCl, 500 mM imidazole
Dialisys buffer 1	50 mM TRIS pH 7.5, 300 mM NaCl, 2 mM ß-ME
Dialisys buffer 2	50 mM TRIS ph 7.5, 300 mM NaCl

Purification of 6 x (His)-hSMUG (25-270) was performed in several steps including affinity purification with Tallon beads as a first step followed by the dialysis and TEV protease treatment, and as the last step a second affinity chromatography using the Äkta start purification system and HiTrap TALLON® Crude 1 ml column. Lysis buffer was added to the pellet from previous step (7 ml per gram of pellet) along with one tablet of Complete EDTA-free protease inhibitor cocktail (Roche) and frozen at -20°C. After the lysis buffer was thawed, the bacterial lysate was supplemented with lysozyme (final concentration 100 µg/ml), DNAse I (final concentration 5 µg/ml), RNAse A (final concentration 5 µg/ml), Tergitol (final concentration 0.5%) and $MgCl_2$ (final concentration 0.5%) followed by 30 min incubation at RT with gentle orbital shaking. The lysate was then sonicated with pulse-on for 10 sec and pulse-off 10 sec using 30% amplitude (Branson Digital Sonifier®) for a total of 30 sec sonication. Afterwards, the insoluble debris were removed by centrifugation at 20 000 rpm for 40 min at 4˚C. The supernatant (crude extract) was placed on ice for following batch purification.
TALON®Metal Affinity Resin (2 ml) was prepared in accordance with manufacturer instructions (TaKaRa) and equilibrated with the equilibration buffer. After equilibration, beads were incubated with 35 ml of crude extract from previous step for 30 min at 4°C. Then the flow through (crude extract after treatment with beads) was separated by centrifugation (900 ×g, 10 min, 4°C) and removed to be analyzed by SDS-PAGE to evaluate the efficiency of 6-(His)-tag binding to the beads. Next, the beads were washed with washing buffer (10 ml) for 10 min, and centrifuged (700 x g, 5 min, 4°C) prior to elution. Elution of 6 × (His)-hSMUG (25-270) was performed using elution buffer 1 (2 ml) incubated with the beads for 10 min at 4°C. Elution fraction containing the protein of interest was collected by centrifugation (700 × g, 5 min, 4°C). The elusion step was repeated once. The fractions from two elusion steps were analyzed by SDS-PAGE as described in section 2.3.5 before dialysis.

Dialysis:
The dialysis was needed to treat the samples with the TEV protease, remove the traces of imidazole and remove the histag for the TEV treated fractions. First elution fractions (2 ml) from previous step were split in two. One was treated with TEV protease (25 µl of AcTEV protease, Themo Fisher Scientific, 12575015) to remove the 6 × (His)-tag, and the other was left untreated. Two fractions (1 ml each) were dialyzed using the Pre-wetted RC tubing MWCO 15 kDa (SpectraPor®) in 2 L dialysis buffer 1 and incubated with gentle stirring o/n at 4°C. For the fraction without TEV a second dialysis using dialysis buffer 2 was performed for 3 h at 4°C to remove the traces of βME which interferes with the BCA protein measurement assay. In order to separate TEV protease, 6 × (His)-peptide and un-cleaved 6 × (His)-hSMUG (25-270) and other contaminants, the second affinity chromatography was performed using HiTrap TALON® crude 1 ml column and Äkta Start purification system. The Äkta Start purification system was prepared according to the manual by washing pumps, valves, fractionation tubes and the column first with water, and then with the equilibration buffer. The 1 ml fraction from dialysis after TEV protease treatment was loaded to the Hitrap Tallon crude 1 ml column using sample valve with the flow rate 1 ml/min. The flow through contacting the protein with cleaved 6 × (His)-tag was collected in two 2 ml fractions. Elutions of the column were made stepwise with 50 mM, 125 mM, 250 mM, 375 mM and 500 mM elution buffer 2, and fractions were collected with 1 ml collection size with the flow rate 1 ml/min. Fractions from flow through were analyzed by nanodrop and BCA.

2.3.5 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis

Buffers
Buffers are listed in Table 9

Table 9: Buffers used for SDS-PAGE analysis

Buffer	Composition
1 × TRIS-Glycine running buffer	25 mM trisma base, 0.192 M glycine, 1% SDS
formamide gel loading buffer	80% formamide, 1 mM EDTA and 1% (w/v) blue dextran

Samples (10 µl) were mixed with formamide gel loading buffer (10 µl) and heated for 5 min at 95°C. The samples were then cooled for two minutes before they were loaded (15 µl) into the wells of a precast SDS-PAGE Gel (BIO-RAD Mini ProteanR TGX™ Precast Gel 10%, 12 well). The 10 µL of Precision plus protein ™ dual color standard (BioRad) was used as a molecular weight standard. The system was run with 1 × Tris-Glycine running buffer for SDS-PAGE, for 30 min at 220 V. The gel was then stained with the Simple Blue Protein Stain (Novex by Life Technologies Simply Blue™ SafeStain) according to the manufacturers manual. After staining, the gel was analyzed on Chemidoc Tough (BioRad). The figure was modified using the Image Lab software (BioRad).

2.3.6 Determination of protein concentration

Protein concentration from previous steps were measured using the BCA Protein Assay Kit (Pierce™) according to the user manual. Briefly, the microplate procedure with a sample to Working Reagent (WR) ratio of 1:8 was employed using the Culture plates (VWR® Tissue Culture Plates, 96 wells-F, sterile). The 9 standards of BSA were prepared in accordance with manufacturers manual (see app. A.3) using equilibration buffer as diluent. The WR was prepared as described in the maufacturers manual. For each standard and unknown sample, 25 µl was transferred to a microplate well along with 200 µl of WR, followed by 30 min incubation at 37°C. Plate was cooled to RT before absorbance was measured at 562 nm on a plate reader (Molecular Devices, SpectraMax® Paradigm® Multi-Mode

Detection Platform). Results from the reading provided a standard curve from where the sample concentrations were calculated.

2.3.7 Verification of protein

Materials

- Primary rabbit anti-SMUG (Abcam, Ab192240)
- Goat anti rabbit- IgG-HRP (SouthernBiotech)

A Western Blot (WB) was performed to verify the protein of interest and confirm the efficiency of TEV cleavage. 10 ng, 20 ng, 50 ng, and 100 ng protein samples with HISTag and without HISTag were used for WB. The samples for SDS-PAGE were treated as above mentioned. For blotting a 0.2 µm PVDF membrane was used as part of the Trans-Blot® Turbo™ Transfer Pack and performed according to the user manual. A Trans Blot® Turbo™ Transfer System was used to blot the gel for 3 minutes. The membrane was blocked by 5% (0,1% Twin 20) for 1 h at 4°C. Then the membrane was incubated with primary rabbit anti-SMUG Ab 1:2000 (Abcam, Ab192240) in 5% PBST milk at 4°C overnight. The blot was then washed 3 × 5 min with PBST (0.1% Tween 20), before being incubated with secondary Goat anti rabbit- IgG-HRP ab 1:2000 (SouthernBiotech) in 5% PBST milk for 1.5 h at RT. Blot was again washed 3 × 5 min and developed using the SuperSignal ™ West Pico Kit (Thermo Scientific). Blot was then analysed using Chemidoc Tough from BioRad.

3 Results

3.1 Effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine in *E. coli.*

Rifampicin is a potent and broad spectrum antibiotic, primarily used to treat tuberculosis in the medicine of today. It works by obstructing transcription by binding to the β-subunit of the RNA polymerase, which is encoded by the *rpoB*$^{+}$ gene in *E. coli.* A multitude of mutations, including all types of base substitutions in this gene, will give rise to rifampicin resistance (Rif^{R}) [62]. In this study resistance to rifampicin was used to measure the mutation rate without and with f^{5}dU addition to a bacterial culture of 2 ml grown for about two days. The total number of bacteria was determined by different dilutions of the final culture plated on glucose dishes, while the number of mutants arisen was determined by transferring undiluted culture on plates with rifampicin added. From each experiment, one mutant colony was harvested and part of their *rpoB*$^{+}$ gene (Rif^{R} region) was amplified and sequenced followed by comparison to the wild type sequence.

3.1.1 *uvrB*$^{-}$

In total 22 assays were started in 2019 which yielded 51 colonies that were chosen for further analysis, in addition 37 mutant bacterial cultures from 2016 were processed and analyzed. Of the 51 analyzed colonies from 2019, five had mutations that fell outside of the sequenced area, and three yielded sequences of too poor quality to make a definitive decision as to which mutation had occurred. Of the 37 analyzed cultures from 2016, eight were of too poor quality to determine the mutations. In addition, three sequences from 2019 and two sequences from 2016 did not give any result in matching sequences when run through the BLAST program of ncbi. That leaves 67 out of the 88 mutants that were successfully analyzed and are presented in Table 10 and Figure 14.

Table 10: Distribution of mutations detected in *Escherichia coli uvrB*$^{-}$ (AB 1885)

Mutation	Spontaneous		0.1mM f^{5}dU	
	%	No.	%	No.
AT→CG	0	0	0	0
GC→AT	54.2	19	71.9	23
GC→CG	5.7	2	3.1	1
GC→TA	34.3	12	15.6	5
AT→TA	2.9	1	6.3	2
AT→GC	2.9	1	3.1	1
TOT	100	35	100	32

The predominant base substitution in both spontaneously arisen and 0.1 mM f^{5}dU induced mutations, is the GC→AT transition, constituting 54.2% and 71.9% of all detected mutations, respectively. The second most common substitution is the GC→TA transversion constituting 34.3% of all spontaneously arisen mutations and 15.6% of all 0.1 mM f^{5}dU induced mutations.
The order of spontaneously arisen mutation distribution is:
GC→AT > GC→TA > GC→CG > AT→TA = AT→GC
The order of 0.1 mM f^{5}dU induced mutation distribution is:
GC→AT > GC→TA > AT→TA > GC→CG = AT→GC
No AT→CG mutations were detected in the strain.

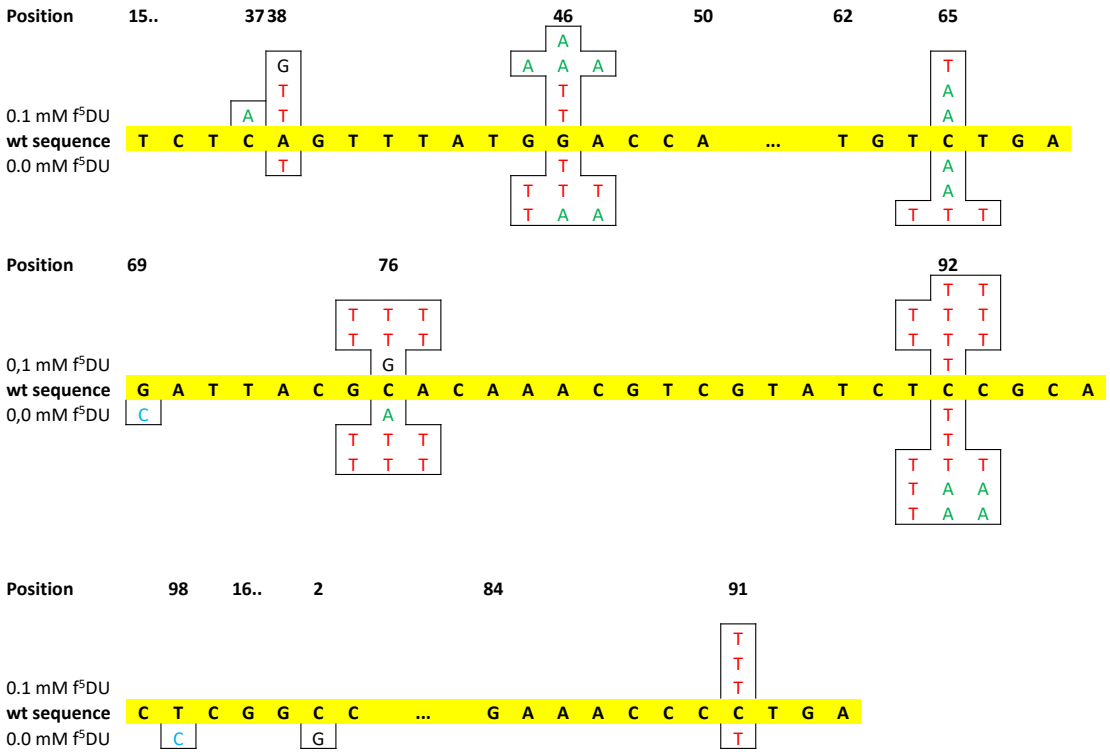


Figure 14: Distribution of detected mutations in the Rif^R region of the *rpoB*$^+$ gene in *Escherichia coli uvrB*$^-$ strain (AB 1885).
Yellow line indicates the wild type sequence. Letters placed above the wild type sequence indicates 0.1 mM f^{5}dU induced mutations while letters placed below the wild type sequence indicates spontaneous mutations. Numbering of the bases is based upon [62].

The mutations detected were distributed between 10 sites within the sequenced part of the *rpoB*$^+$ gene, where six of the sites were shared between spontaneously arisen mutations and mutations induced by the addition of 0.1 mM f^{5}dU. Three sites were exclusive to spontaneous mutation and one was exclusive to a 0.1 mM f^{5}dU induced mutation.

3.1.2 *uvrC*$^-$

Only two assays were started with the *uvrC*$^-$ (AB1884) strain, and 9 mutant colonies were further analyzed. One sequence was of too poor quality to determine the actual mutation, and the results of the eight remaining sequences are presented in Table 11 and Figure 15.

Table 11: Distribution of mutations detected in *Escherichia coli* AB 1884

Mutation	Spontaneous		0.1mM f^{5}dU	
	%	No.	%	No.
AT→CG	0	0	0	0
GC→AT	100	2	50	3
GC→CG	0	0	0	0
GC→TA	0	0	33.3	2
AT→TA	0	0	0	0
AT→GC	0	0	16.7	1
TOT	100	2	100	6

The results from the *uvrC*⁻ mutagenesis experiments provide little information unless combined with previously obtained data, which is undertaken in the Discussion section.

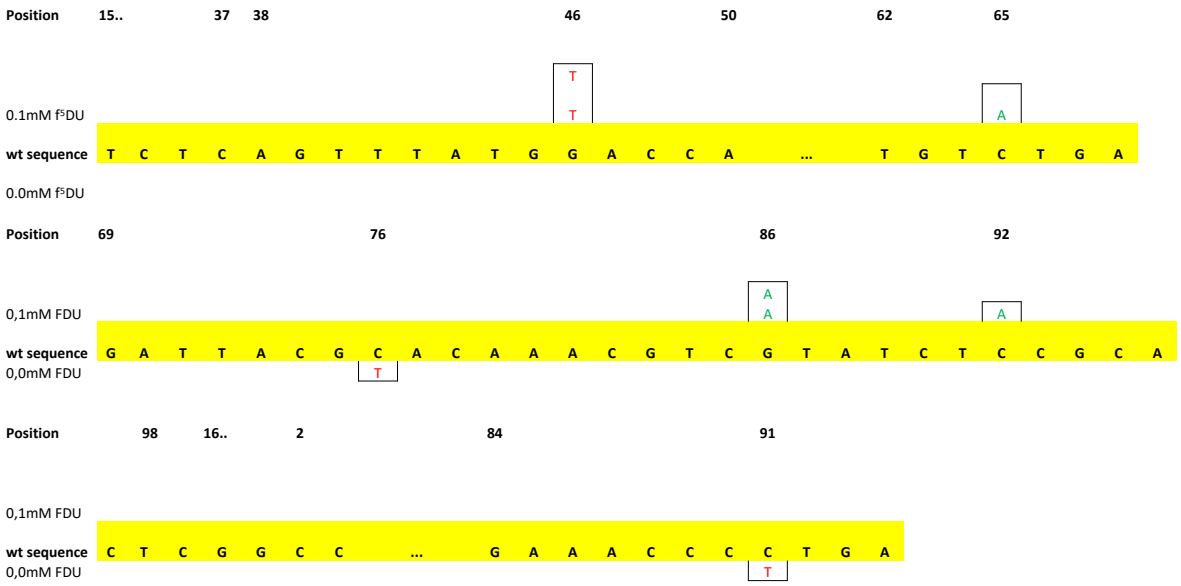


Figure 15: Distribution of detected mutations in the Rif^R region of the *rpoB* gene in *Escherichia coli* AB 1884 strain. Yellow line indicates the wild type sequence. Letters placed above the wt sequence indicates 0.1 mM f^{5}dU induced mutations while letters placed below the wt sequence indicates spontaneous mutations. Numbering of the bases is based upon [62].

3.2 DNA glycosylase activities for N^4,5-dimethylcytosine

The method used to determine DNA glycosylase activity in this thesis relied upon a method where fluorescently tagged oligodeoxyribonucleotides with $m^{N4,5}C$ inserted at a specific position were hybridized with complimentary oligodeoxyribonucleotides with A, C, G or T placed opposite the damaged base. As all tested enzymes were monofunctional DNA glycosylases, the samples were treated with NaOH and heat to cleave the resultant AP site followed by denaturing PAGE to separate cleaved product DNA from un-cleaved substrate DNA. As positive control for enzyme activity, a verified control substrate was submitted to the same conditions as the tested substrates and run on the same gel (lane 10 in all gels).

For each enzyme, three parallels were performed to confirm the results. None of the enzymes tested showed any activity against the $m^{N4,5}C$ lesion when placed opposite any of the four native DNA bases at the tested concentrations. Results are presented in figures 16-18, and additional gels are included in the appendices.

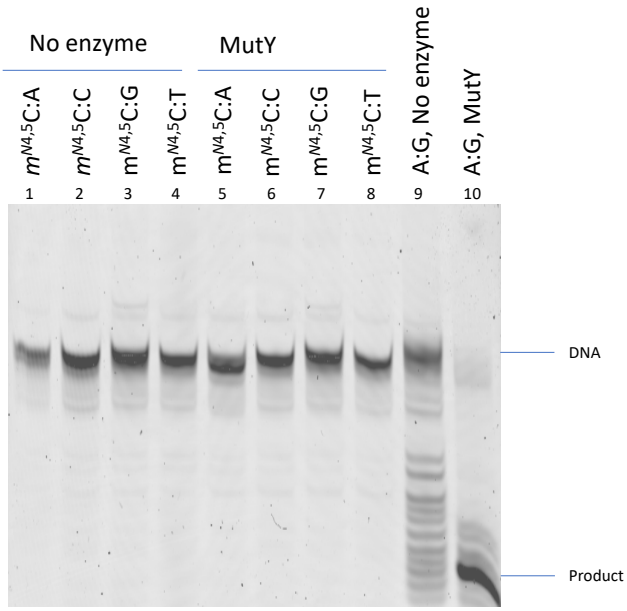


Figure 16: Denaturing PAGE gel of MutY shows no activity towards $m^{N4,5}C$ when placed opposite A, C, G or T
Lanes 1-4 are negative controls with no added enzyme to the substrates, lanes 5-8 are substrates with MutY enzyme added, lane 9 is negative control of control substrate (A:G) with no enzyme added and lane 10 is the positive control for enzyme activity. Incubation was 1 h at 37°C.

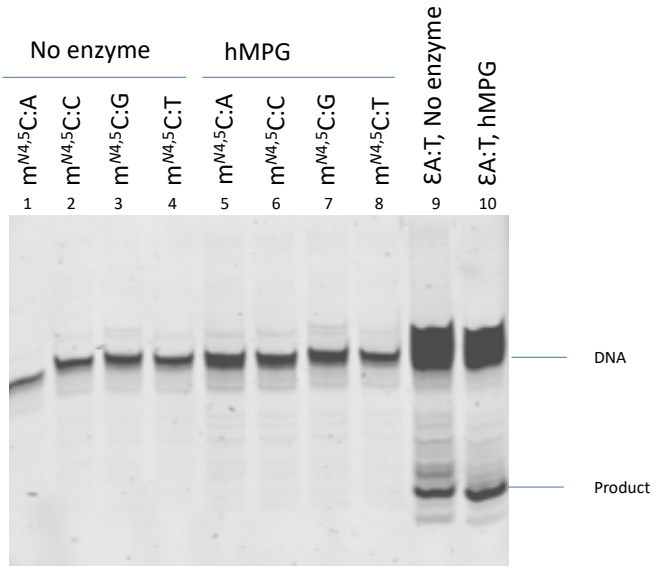


Figure 17: Denaturing PAGE gel of hMPG shows no activity towards $m^{N4,5}C$ when placed opposite A, C, G or T
Lanes 1-4 are negative controls with no added enzyme to the substrates, lanes 5-8 are substrates with MPG enzyme added, lane 9 is negative control of control substrate (εA:T) with no added enzyme and lane 10 is the positive control for enzyme activity. Incubation was 1 h at 37°C.

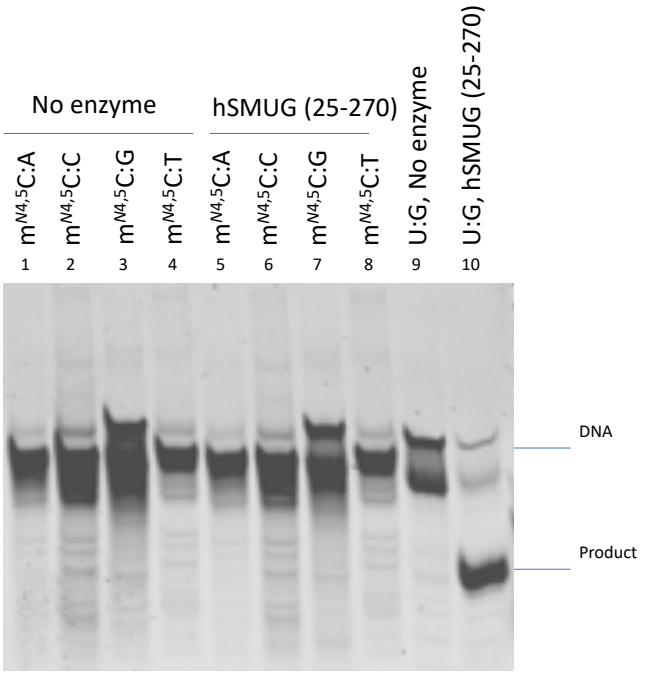


Figure 18: Denaturing PAGE gel of hSMUG (25-270) shows no activity towards $m^{N4,5}C$ when placed opposite A, C, G or T
Lanes 1-4 are negative controls with no added enzyme to the substrates, lanes 5-8 are substrates with hSMUG (25-270) enzyme added, lane 9 is negative control of control substrate (U:G) with no added enzyme and lane 10 is the positive control for enzyme activity. Incubation was 1 h at 37°C.

3.3 Production and purification of hSMUG (25-270)

3.3.1 Competent cells and transformation

The competent cells were successfully transferred with the desired plasmid, confirmed by transformation by the control plasmid PUC 19.

3.3.2 Autoinduction

The transfer of one single colony containing the desired plasmid into auto induction media resulted in 5 g of cell pellet with expressed protein per 500 ml of media.

3.3.3 Purification

After the first batch purification the samples were analyzed by SDS-PAGE, and the results are presented in Figure 19. The protein bands are clearly visible, with the expected decrease in intensity after the second elution.

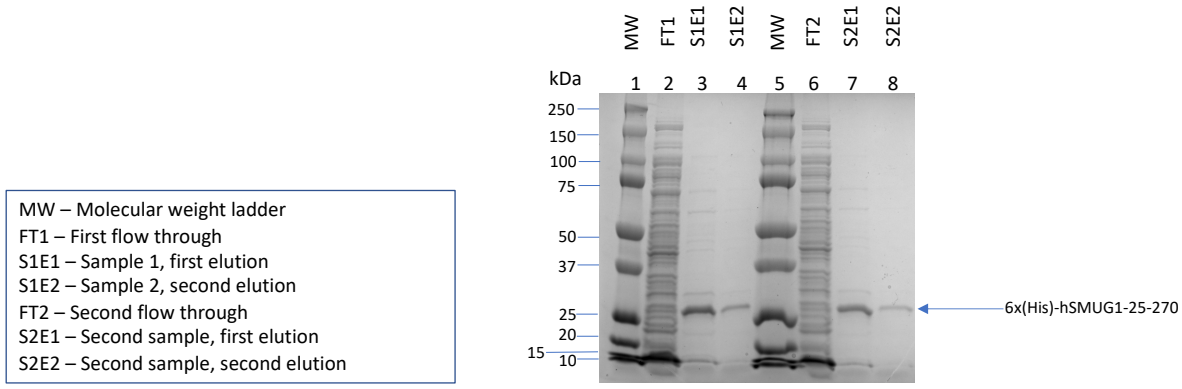


Figure 19: SDS-PAGE gel of eluted samples of purified hSMUG (25-270)

After the dialisys the next step was a second affinity chromatography using the Äkta Start purification system. Data collected through this analysis is presented in Figure 21, where the inserted percentages indicate amount of imidazole used to elute the sample.

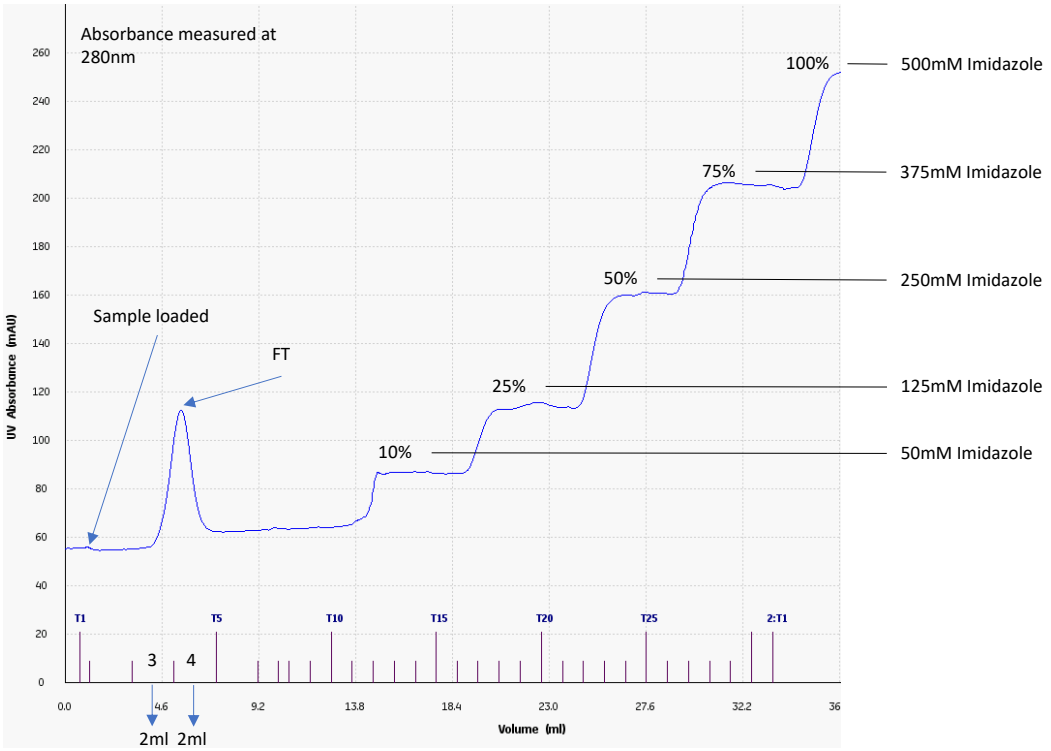


Figure 20: Chromatogram from the Äkta Start purification system

The first peak of the chromatogram in Figure 20 was collected and used to perform SDS-PAGE analysis, followed by a western blot which is presented in figure 21.

3.3.4 Determining concentration

Concentration was determined by BCA protein assay, followed by microplate analysis and the results were compared to the standard curve and found to be 96 µg/ml (the OD measurements are presented in Table 12).

Table 12: OD measurements from BCA protein assay and subsequent microplate analysis

Samples	Wells	OD values	Concentration (µg/ml)	Mean concentration	SD	CV
1	A11	0,148	122,729	120,297	2,591	2,2
	B11	0,142	117,572			
	C11	0,145	120,588			
2	A12	0,070	47,415	44,334	4,442	10,0
	B12	0,062	39,241			
	C12	0,069	46,345			
3	D1	0,056	33,396	22,310	11,094	49,7
	E1	0,033	11,120			
	F1	0,045	22,505			
4	D2	0,018	-3,670	-12,558	9,962	79,3
	E2	-0,002	-23,326			
	F2	0,011	-10,676			

3.3.5 Protein verification

Protein was verified by a Western Blot, which also established the difference in size of the 6x(His)Tag bound protein and the cleaved protein as shown in Figure 21.

Lane 1 – 10 ng sample
Lane 2 – 20 ng sample
Lane 3 – 50 ng sample
Lane 4 – 100 ng sample
Lane 5 – MW ladder (+overflow from lane 4)
Lane 6 – Overflow from lane 4
Lane 7 – MW ladder
Lane 8 – 10 ng sample
Lane 9 – 20 ng sample
Lane 10 – 50 ng sample
Lane 11 – 100 ng sample

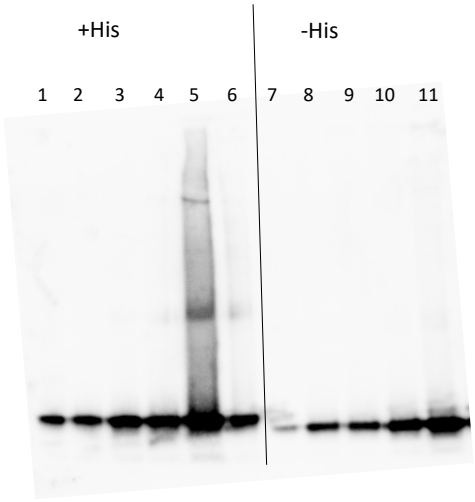


Figure 21: Western blot showing the size difference of the his tag bound protein and cleaved protein.

4 Discussion

4.1 Effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine in *E. coli.*

This thesis work combined with previously obtained results from the project provide an experimentally solid basis to draw conclusions. Therefore, the first part of this section is dedicated to presenting the combined findings of the study.

Mutation frequencies and distribution

In all figures regarding distribution of mutations, the position (base numbering) is based upon the start codon, ATG, of the *rpoB*$^{+}$ gene with the A in the ATG codon as base number one [62].

Table 13: Percentage and distribution of base substitutions among Rif^{R} mutants arisen spontaneously and induced by the addition of 0.1 mM f^{5}dU to exponentially growing cells of *E. coli*.

	Wild type*		***uvrA*$^{-}$***		***uvrB*$^{-}$**		***uvrC*$^{-}$**	
f^{5}dU	0 mM	0.1 mM	0 mM	0.1 mM	0 mM	0.1 mM	0 mM	0.1 mM
	%	%	%	%	%	%	%	%
AT→CG	11.6 (10)	2.9 (2)	1.6 (1)	1.2 (1)	0	0	0	0
GC→AT	34.9 (30)	11.8 (8)	61.9 (39)	38.2 (34)	50 (26)	61.3 (38)	65.6 (40)	52.5 (32)
GC→CG	3.5 (3)	0	0	1.2 (1)	7.7 (4)	1.6 (1)	0	0
GC→TA	13.9 (12)	4.4 (3)	25.4 (16)	22.5 (20)	30.8 (16)	16.1 (10)	23.0 (14)	36.1 (22)
AT→TA	10.5 (9)	7.4 (5)	4.8 (3)	11.2 (10)	3.8 (2)	6.5 (4)	1.6 (1)	0
AT→GC	13.9 (12)	66.2 (45)	3.2 (2)	14.6 (13)	1.9 (1)	4.8 (3)	1.6 (1)	6.6 (4)
Indels	2.3 (2)	0	0	0	0	0	0	0
Unknown	9.3 (8)	7.4 (5)	3.2 (2)	11.2 (10)	5.8 (3)	9.7 (6)	8.2 (5)	4.9 (3)
TOTAL	100 (86)	100 (68)	100 (63)	100 (89)	100 (52)	100 (62)	100 (61)	100 (61)

Numbers in parenthesis indicates number of mutants

* Data collected from [4]

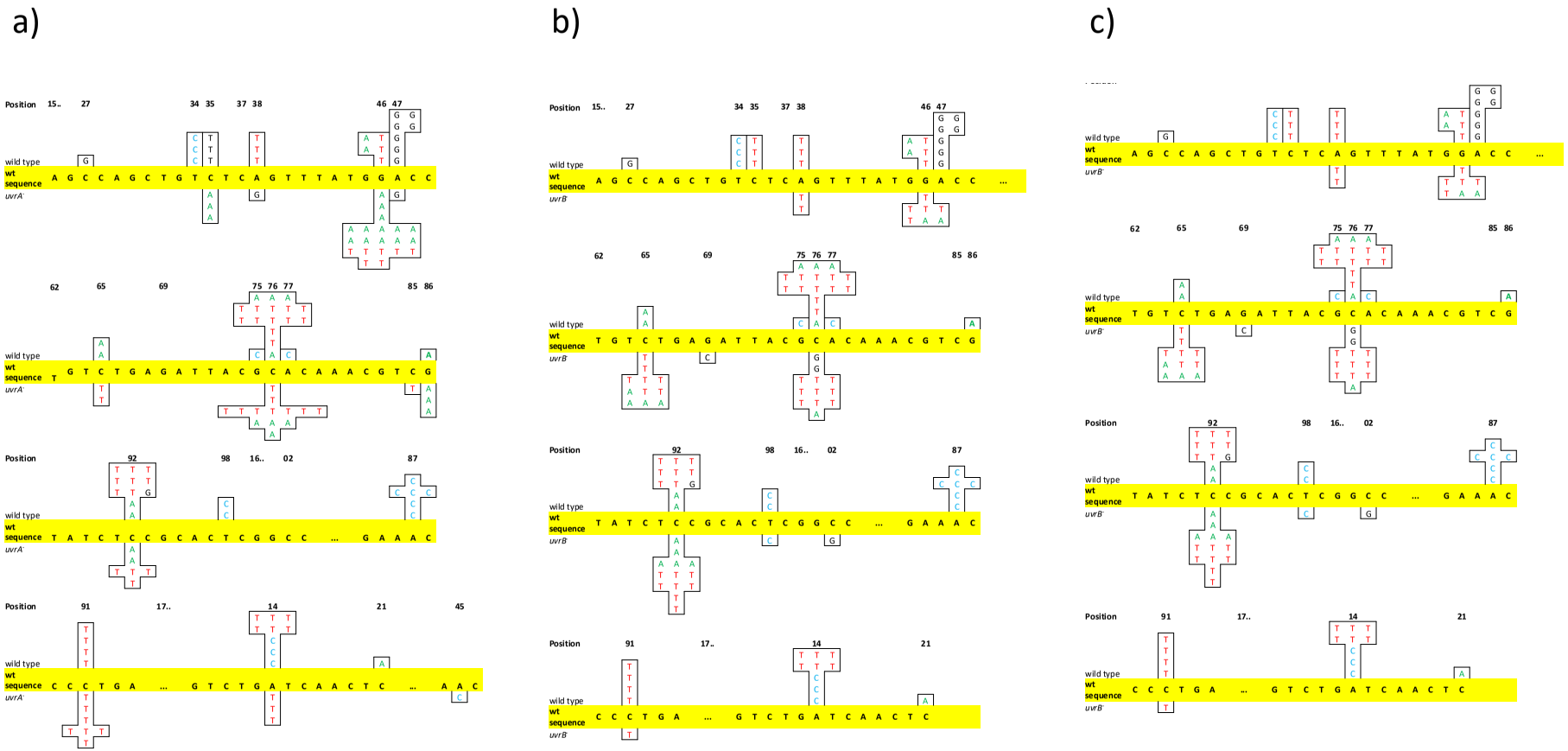


Figure 22: Distribution of spontaneous mutations in all repair deficient strains compared to wild type. Yellow line indicates the wild type sequence of the sequenced part of the *rpoB*⁺ gene.
a) above yellow line: spontaneous wild type mutations, below yellow line: spontaneous *uvrA*⁻ mutations
b) above yellow line: spontaneous wild type mutations, below yellow line: spontaneous *uvrB*⁻ mutations
c: above yellow line: spontaneous wild type mutations, below yellow line: spontaneous *uvrC*⁻ mutations

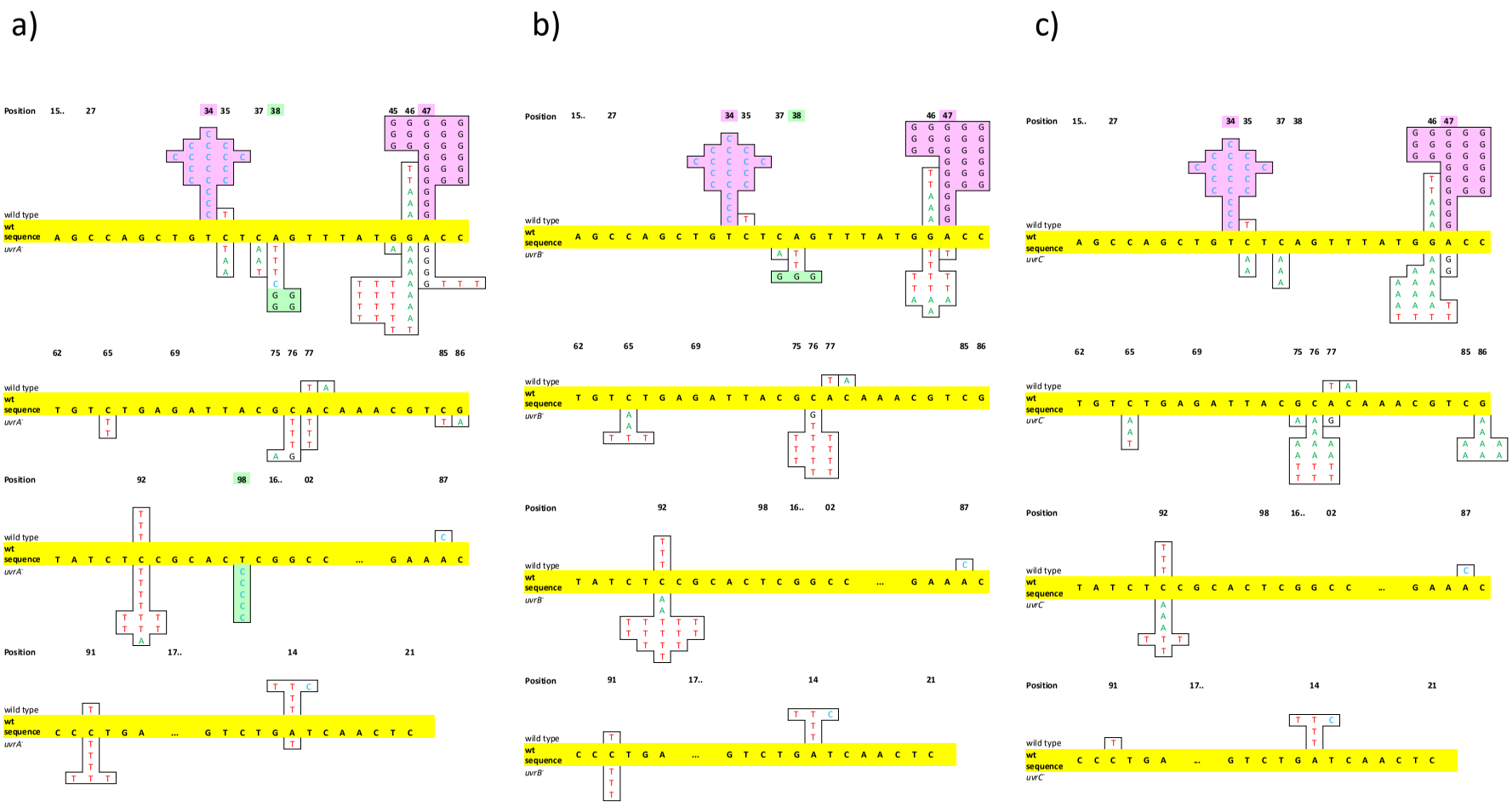


Figure 23: Distribution of f^{5}dU induced mutations in all repair deficient strains compared to wild type. Yellow line indicates the wild type sequence of the sequenced part of the *rpoB*$^+$ gene.
a) above yellow line: f^{5}dU induced wild type mutations, below yellow line: f^{5}dU induced *uvrA*$^-$ mutations
b) above yellow line: f^{5}dU induced wild type mutations, below yellow line: f^{5}dU induced *uvrB*$^-$ mutations
c: above yellow line: f^{5}dU induced wild type mutations, below yellow line: f^{5}dU induced *uvrC*$^-$ mutations

Position		15..	27	34	35	37	38	45	46	47	65
0.0 mM fdU	wt		1.32	3.95	3.95		3.95		6.58	9.21	2.63
	uvrA⁻				4.92		1.64		32.79	1.64	3.28
	uvrB⁻						4.08		14.29		22.45
	uvrC⁻				1.79	3.57			30.36	1.79	3.57
wt sequence		**A**	**G C C A G C T G T C T C A G T T T A T G G A C C A**								**... T G T C T G A**
0.1 mM fdU	wt			28.57	1.59				7.94	42.86	
	uvrA⁻				3.80	3.80	10.13	1.27	27.85	8.86	2.53
	UvrB⁻					1.79	8.93		21.43	1.79	8.93
	UvrC⁻				3.45	5.17			29.31	3.45	5.17

Position		69	75	76	77	78	85	86	92	98	16..	02	04
0.0 mM fdU	wt		1.32	21.05				1.32	14.47	2.63			
	uvrA⁻			21.31			1.64	4.92	9.84				
	uvrB⁻	2.04		24.49					26.53	2.04		2.04	
	uvrC⁻			16.07			1.79	21.43	14.29				
wt sequence		**G A T T A C G C A C A A A C G T C G T A T C T C C G C A C T C G G C C C**											
0.1 mM fdU	wt				1.59	1.59			4.76				
	uvrA⁻			6.33	3.80		1.27	1.27	13.92	6.33			
	UvrB⁻			23.21					28.57				
	UvrC⁻		1.72	24.14	1.72			13.79	12.07				

Position		84	87	91	17..	14	21	45
0.0 mM fdU	wt		7.89	5.26		11.84	1.32	
	uvrA⁻			11.48		4.92		1.64
	uvrB⁻			2.04				
	uvrC⁻			3.57		1.79		
wt sequence		**G A A A C C C C T G A ... G T C T G A T C A A C T C T ... C T A A C**						
0.1 mM fdU	wt		1.59	1.59		7.94		
	uvrA⁻			7.59		1.27		
	UvrB⁻			5.36				
	UvrC⁻							

Figure 24: Map showing mutation distribution within the sequenced part of the *rpoB*⁺ gene of all four investigated *E. Coli* strains, both spontaneous and f⁵dU induced.
Each number represents the percentage of detectable base substitutions within each strain with or without 0.1 mM f⁵dU added to the growth media.

Mutation rates

Since bacteria grow exponentially and mutational events are assumed to be stochastic [63], the information gathered from the mutation frequencies is limited. In short this can be explained by the fact that a mutation that originates early in the growth period prior to selection, will yield more daughter cells carrying that mutation than mutations that arise later in the growth period, a phenomenon referred to as a "jackpot culture", first described by Luria and Delbrück [64]. By calculating the total mutation rates for each individual strain and the mutation rate for each type of mutation, the information provided is more accurate and not skewed by when the mutational event occurred. In short, the mutation rate can be explained as the theoretical probability of a mutational event per cell division.

Mutation rate calculations were based on the p_0 method first described in reference [64].
The results of the diluted cultures plated on glucose dishes was used to determine the number of cells/100 µl at the end of each experiment (N_t), where n = cells/100µl. In addition the number of cells at the start of each experiment was also known (N_0). From these numbers the number of cell divisions (and thus potentially mutagenic events) was calculated.
Mutation rate calculations are based upon a selection of the performed experiments where the n-value was used as a discriminating factor. The interval was set to n= $0.5x10^8$ -$1.5x10^8$ cells/100 µl.
All numbers and calculations are provided in appendix A.1.4.

Table 14: Mutation rates for Rif^R in exponentially growing cells of wild type, *uvrA*$^-$, *uvrB*$^-$ and *uvrC*$^-$ cells

	Wild type		***uvrA*$^-$**		***uvrB*$^-$**		***uvrC*$^-$**	
	Mutation rate ($\times 10^{-9}$)							
f^{5}dU		fold		fold		fold		fold
0.0 mM	1.27	1	0.99	1	1.43	1	2.29	1
0.1 mM	2.4	1.89	1.24	1.25	3.98	2.78	4.19	1.83

Table 15: Mutation rate distribution for Rif^R in exponentially growing cells of wild type, *uvrA*$^-$, *uvrB*$^-$ and *uvrC*$^-$ cells

	Wild type		***uvrA*$^-$**		***uvrB*$^-$**		***uvrC*$^-$**	
	Mutation rates ($\times 10^{-9}$)							
	0.0 mM	0.1 mM	0.0 mM	0.1 mM	0.0 mM	0.1 mM	0.0 mM	0.1 mM
AT→CG	0.15	0.07	0.02	0.01	0	0	0	0
GC→AT	0.44	0.28	0.61	0.47	0.72	2.44	1.50	2.20
GC→CG	0.04	0	0	0.01	0.11	0.06	0	0
GC→TA	0.18	0.11	0.25	0.28	0.44	0.64	0.53	1.51
AT→TA	0.13	0.18	0.05	0.14	0.05	0.26	0.04	0
AT→GC	0.18	1.59	0.03	0.18	0.03	0.19	0.04	0.28

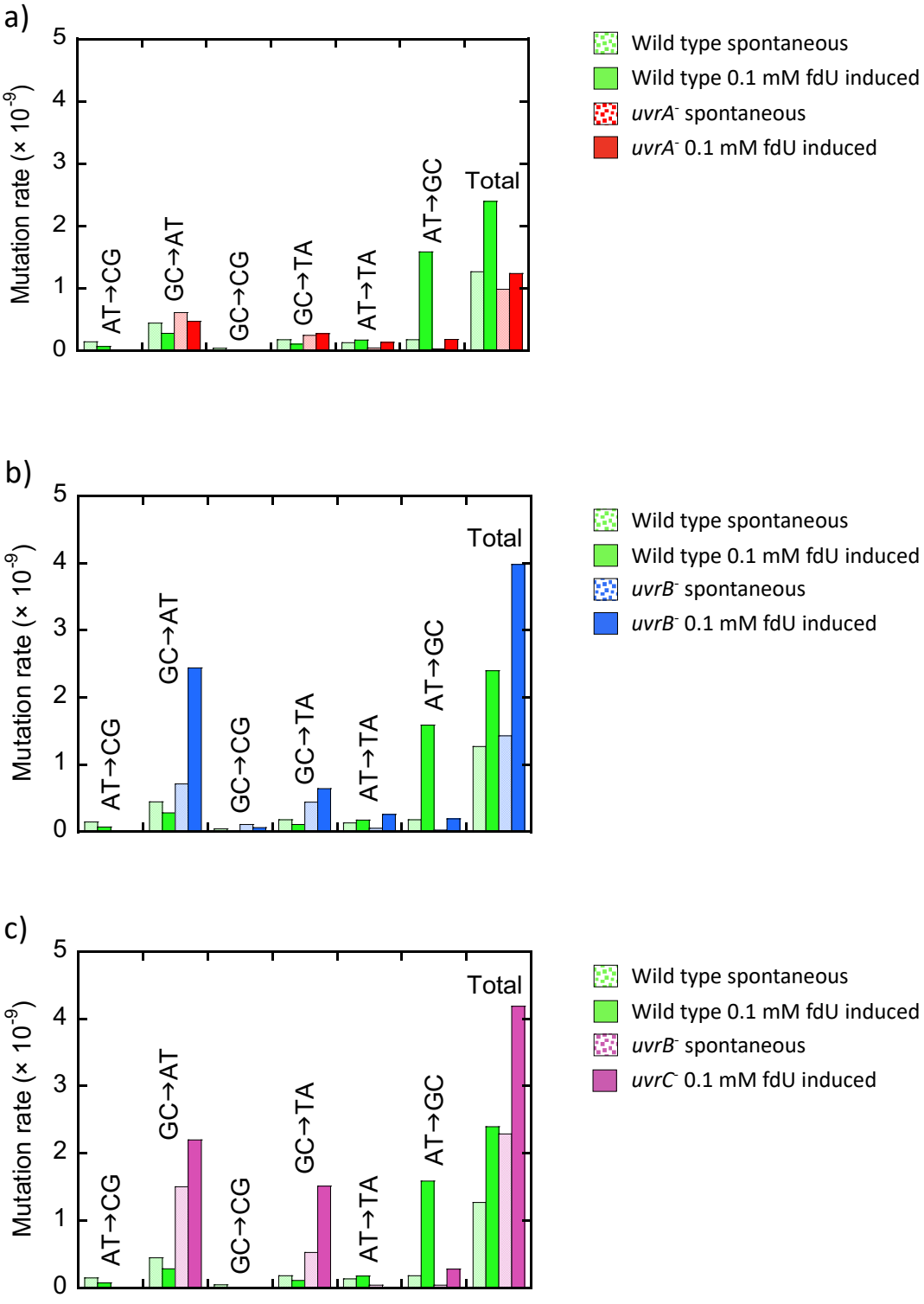


Figure 25: Mutation rate comparison between wild type and the three tested repair deficient strains of *E. Coli*, with or without 0.1 mM f^{5}dU added to the growth medium of exponentially growing cells.
a) Wild type compared to *uvrA*$^-$
b) Wild type compared to *uvrB*$^-$
c) Wild type compared to *uvrC*$^-$

Effects of adding 0.1 mM f^{5}dU to the growth medium of exponentially growing E. coli cells

Early experiments performed on the *uvrB*$^-$ strain showed a decrease in relative growth of 6% when 0.1 mM f^{5}dU was added to the growth medium of exponentially growing cells. This indicates that the additive is moderately cytotoxic, and mainly mutagenic, which is in accordance with previous findings [3].

NER involvement in mutation induction by f^{5}dU in E. coli

There is *in vitro* evidence that the NER pathway can repair f^{5}U lesions in mammalian cells [19]. Previous experiments performed within our group using *uvrA*$^-$ cells indicated that the *uvrA*$^+$ gene was highly involved in promotion and somewhat involved in alleviation of f^{5}dU induced mutations, based upon the change in mutational pattern observed between the wild type and *uvrA*$^-$ strains (unpublished results). This was evident at positions 1534 and 1547, which were clear hot spots for f^{5}dU induced mutations in the wild type. All f^{5}dU induced AT→GC transitions in wild type were found at these two sites. At position 1534 no mutations were detected in the *uvrA*$^-$ strain, and at position 1547 few were detected (Figure 23 a).
To determine if these findings were limited to UvrA involvement or if other players in the NER pathway were also involved, the experiments were started on *uvrB*$^-$ cells and *uvrC*$^-$ cells.
The results show that the same pattern holds true for both strains (Figure 23 b and c), where no or very few mutations were detected at positions 1534 and 1547. This is a clear indication that all the NER proteins, and thus the NER pathway are involved in mutation induction by f^{5}dU.

Even though the findings in this study indicate that the NER pathway is involved in promoting f^{5}dU induced mutations, and by an extension, repair of f^{5}dU induced lesions, the BER pathway is most likely the main repair mechanism for f^{5}U, due to the small distortion to the DNA structure caused by this lesion [3].
AlkA is the primary glycosylase for repairing f^{5}U in *E.coli,* with several back up glycosylases *[51, 65]*,

Addition of 0.1 mM f^{5}dU to the growth medium of exponentially growing E. coli cells increase the overall mutation rate

The total mutation rate did not increase significantly in the *uvrA*$^-$ strain (1.3 fold) as apposed to somewhat less than 2 fold in the *uvrC*$^-$ and 2.8 fold in the *uvrB*$^-$ strains (Table 14). However, the types of mutations contributing to the increase differs between the strains. The insignificant increase in the *uvrA*$^-$ cells will be discussed later, while the focus here will be on the difference between the wt, *uvrB*$^-$ and *uvrC*$^-$ strains.

The main contributor to the increase in mutation rate observed in wt is the AT→GC transition (Figure 25). However, the same cannot be said for *UvrC*$^-$ and *uvrB*$^-$, where the main contributor is the GC→AT transition and the GC→TA transversion, respectively. There is no reasonable explanation as to why the three strains behave so differently when incorporating f^{5}U into DNA (Figure 26). Maybe wild type and *uvrC*$^-$ seem to favor the incorporation of the keto form o f^{5}U, while *uvrB*$^-$ favor the ionized form?

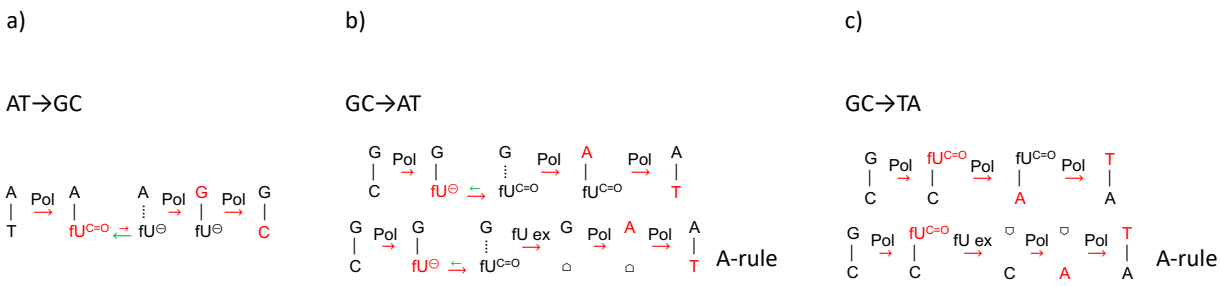


Figure 26: Proposed models for mutation induction by f^{5}dU.

A possibly unknown function of UvrA

All strains with an intact *uvrA*$^+$ gene show a distinct increase in the overall mutation rate when exposed to 0.1 mM f^{5}dU (Figure 26 b and c), whereas the mutation rate in the *uvrA*$^-$ strain only has a very slight or no increase (Figure 26 a). The NER pathway is disabled in all three repair deficient strains, yet the findings in this study suggest that all NER proteins are involved in mutation induction by f^{5}dU induced lesions. Why then, does the *uvrA*$^-$ strain show so little increase in the total mutation rate compared to the other strains?

As discussed in the introduction, UvrA initiates the NER pathway by facilitating binding between UvrB and dsDNA before dissociating, leaving behind a stable UvrB-DNA complex which then recruits UvrC to perform the incisions. UvrA is not involved in the actual repair, it could therefore be regarded as a facilitator more than a repair protein. Several studies have proposed UvrA as a "molecular matchmaker" [27, 66], which utilize the energy released from ATP hydrolysis to perform a conformational change within the DNA binding protein (see Introduction). UvrA fulfills all the proposed criteria for being a molecular matchmaker [67].

1: Affinity of the two molecules in question must be low when the matchmaker is not present.
-UvrB will not bind dsDNA in solution, only ssDNA [31].
2: Stable complex formation between the two molecules must be obtained.
-Studies have shown that the half-life of the UvrB-DNA complex is 2-3 hours [68].
3: Conformational change is required without covalent interactions.
-The cryptic ATPase activity of UvrB is codependent on the presence of both UvrA and DNA, which is explained by the conformational change within the UvrB protein upon binding to UvrA and DNA [30].
4: The matchmaker itself or one of the matched molecules must be an ATPase.
- UvrA is a DNA-independent ATPase, and the UvrB protein exhibits cryptic ATPase activity when bound to UvrA and DNA[30].
5: The matchmaker must leave the complex after stable complex formation is achieved.
-Upon damage detection and verification, UvrA will facilitate binding of UvrB to DNA through ATP hydrolysis. UvrA will then dissociate, leaving behind a UvrB-DNA complex which has a high affinity for UvrC [26].

All strains where UvrA is present show a distinct increase in total mutation rate, this could be explained if UvrA exhibits more functions than is presently known. For instance, if UvrA has a function where it recruits other repair enzymes upon detecting f^{5}U lesions in addition to UvrB.

4.2 DNA glycosylase activities for N^4,5-dimethylcytosine

The results from this part of the project is also combined with earlier results within the same project. As there was no activity detected in any of the investigated DNA glycosylases, we are close to conclude that Nei is the major glycosylase for repair of $m^{N4,5}C$ in DNA (highest activity opposite G), and together with Fpg provide repair of the lesion opposite all common bases in DNA [17]. To reach a final conclusion it is needed to investigate 3-methyladenine-DNA-glycosylase I (Tag), which, however, is unlikely to exhibit activity for $m^{N4,5}C$ due to its very narrow substrate specificity [55]

Table 16: DNA glycosylase activity against $m^{N4,5}C$ paired with normal DNA bases in short oligonucleotides

DNA pairing/ glycosylases		$m^{N4,5}C$:G	$m^{N4,5}C$:C	$m^{N4,5}C$:A	$m^{N4,5}C$:T
E. coli	Fpg	-	+++	-	+
	Nei	+++	-	+	+
	Nth	-	-	-	-
	MutY*	-	-	-	-
	Ung	-	-	-	-
	Mug	-	-	-	-
Human	hOGG1	-	-	-	-
	hSMUG1	-	-	-	-
	hTDG	-	-	-	-
	hNEIL1	-	-	-	-
	hNEIL2	-	-	-	-
	hNEIL3	-	-	-	-
	hMPG*	-	-	-	-
	hUNG	-	-	-	-

High activity: +++, low activity: +, no activity: -
*Tested as part of this master thesis

Adapted from [17]

4.3 Future prospectives

Some questions were answered as a result of this thesis work, yet more were raised, as is often the case when working in science. Questions that were raised include why the different strains seem to incorporate f^5dU differently leading to distinct raise in specific mutations, and maybe most importantly, why does *uvrA*$^-$ not experience the same rise in total mutation rates as other investigated strains? Could there be an unknown function of the UvrA protein not yet discovered? Also, the effects of higher f^5dU concentrations such as 0.2 mM for *uvrA*$^-$, *uvrB*$^-$ and *uvrC*$^-$ strains should be investigated. In addition, an investigation to see the effects of f^5dU on the double mutant *uvrA*$^-$ *alkA*$^-$ could shed some more insight to the repair of f^5U in DNA. Another important issue to consider is whether or not the SOS response is induced as a consequence of the addition of f^5dU to the growth medium.

References

Uncategorized References

1. Krwawicz, J., et al., *Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease.* Acta Biochim Pol, 2007. **54**(3): p. 413-34.
2. Cadet, J. and J.R. Wagner, *DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation.* Cold Spring Harb Perspect Biol, 2013. **5**(2).
3. Bjelland, S. and E. Seeberg, *Mutagenicity, toxicity and repair of DNA base damage induced by oxidation.* Mutat Res, 2003. **531**(1-2): p. 37-80.
4. Grøsvik, K., et al., *The Escherichia coli alkA gene is activated to alleviate mutagenesis by an oxidised base.*
5. Ånensen, H., et al., *Mutations induced by 5-formyl-2'-deoxyuridine in Escherichia coli include base substitutions that can arise from mispairs of 5-formyluracil with guanine, cytosine and thymine.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2001. **476**(1-2, 9 May 2001): p. 99-107.
6. Sedgwick, B., *Repairing DNA-methylation damage.* Nat Rev Mol Cell Biol, 2004. **5**(2): p. 148-57.
7. Breiling, A. and F. Lyko, *Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond.* Epigenetics Chromatin, 2015. **8**: p. 24.
8. Chen, Q.W., et al., *Epigenetic regulation and cancer (review).* Oncol Rep, 2014. **31**(2): p. 523-32.
9. Rydberg, B. and T. Lindahl, *Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction.* EMBO J, 1982. **1**(2): p. 211-6.
10. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.
11. Klungland, A. and A.B. Robertson, *Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine.* Free Radic Biol Med, 2017. **107**: p. 62-68.
12. Cooper, D.N. and H. Youssoufian, *The CpG dinucleotide and human genetic disease.* Hum Genet, 1988. **78**(2): p. 151-5.
13. Ratel, D., et al., *N6-methyladenine: the other methylated base of DNA.* Bioessays, 2006. **28**(3): p. 309-15.
14. Yao, B., et al., *DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress.* Nat Commun, 2017. **8**(1): p. 1122.
15. Yao, B., et al., *Active N(6)-Methyladenine Demethylation by DMAD Regulates Gene Expression by Coordinating with Polycomb Protein in Neurons.* Mol Cell, 2018. **71**(5): p. 848-857 e6.
16. Klimasauskas, S., et al., *N4,5-dimethylcytosine, a novel hypermodified base in DNA.* Nucleic Acids Res Suppl, 2002(2): p. 73-4.
17. Alexeeva, M., et al., *Excision of the doubly methylated base N(4),5-dimethylcytosine from DNA by Escherichia coli Nei and Fpg proteins.* Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1748).
18. Ames, B.N., M.K. Shigenaga, and T.M. Hagen, *Oxidants, antioxidants, and the degenerative diseases of aging.* Proc Natl Acad Sci U S A, 1993. **90**(17): p. 7915-22.
19. Kino, K., et al., *Nucleotide excision repair of 5-formyluracil in vitro is enhanced by the presence of mismatched bases.* Biochemistry, 2004. **43**(10): p. 2682-7.
20. Muruzabal-Lecumberri, I., et al., *Alleviation and promotion of damage-specific mutation induction in Escherichia coli are highly dependent on the uvrA gene.*

21. Wirth, N., et al., *Conservation and Divergence in Nucleotide Excision Repair Lesion Recognition.* J Biol Chem, 2016. **291**(36): p. 18932-46.
22. Verhoeven, E.E., et al., *The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands.* EMBO J, 2002. **21**(15): p. 4196-205.
23. Malta, E., G.F. Moolenaar, and N. Goosen, *Dynamics of the UvrABC nucleotide excision repair proteins analyzed by fluorescence resonance energy transfer.* Biochemistry, 2007. **46**(31): p. 9080-8.
24. Pakotiprapha, D., et al., *Structure and mechanism of the UvrA-UvrB DNA damage sensor.* Nat Struct Mol Biol, 2012. **19**(3): p. 291-8.
25. Thiagalingam, S. and L. Grossman, *Both ATPase sites of Escherichia coli UvrA have functional roles in nucleotide excision repair.* J Biol Chem, 1991. **266**(17): p. 11395-403.
26. Friedberg, E.C., et al., *DNA Repair and mutagenesis*. 2 ed. 2006, Washington, DC 20036-2904: ASM Press.
27. Sancar, A. and M.S. Tang, *Nucleotide excision repair.* Photochem Photobiol, 1993. **57**(5): p. 905-21.
28. Myles, G.M. and A. Sancar, *Isolation and characterization of functional domains of UvrA.* Biochemistry, 1991. **30**(16): p. 3834-40.
29. Timmins, J., et al., *Structural and mutational analyses of Deinococcus radiodurans UvrA2 provide insight into DNA binding and damage recognition by UvrAs.* Structure, 2009. **17**(4): p. 547-58.
30. Caron, P.R. and L. Grossman, *Involvement of a cryptic ATPase activity of UvrB and its proteolysis product, UvrB* in DNA repair.* Nucleic Acids Res, 1988. **16**(22): p. 10891-902.
31. Van Houten, B., et al., *DNase I footprint of ABC excinuclease.* J Biol Chem, 1987. **262**(27): p. 13180-7.
32. Moolenaar, G.F., L. Hoglund, and N. Goosen, *Clue to damage recognition by UvrB: residues in the beta-hairpin structure prevent binding to non-damaged DNA.* EMBO J, 2001. **20**(21): p. 6140-9.
33. Hildebrand, E.L. and L. Grossman, *Oligomerization of the UvrB nucleotide excision repair protein of Escherichia coli.* J Biol Chem, 1999. **274**(39): p. 27885-90.
34. Moolenaar, G.F., et al., *Function of the homologous regions of the Escherichia coli DNA excision repair proteins UvrB and UvrC in stabilization of the UvrBC-DNA complex and in 3'-incision.* Mutat Res, 1997. **385**(3): p. 195-203.
35. Zou, Y., et al., *Formation of DNA repair intermediates and incision by the ATP-dependent UvrB-UvrC endonuclease.* J Biol Chem, 1997. **272**(8): p. 4820-7.
36. Moolenaar, G.F., et al., *Characterization of the Escherichia coli damage-independent UvrBC endonuclease activity.* J Biol Chem, 1998. **273**(52): p. 34896-903.
37. Singh, S., et al., *Solution structure and DNA-binding properties of the C-terminal domain of UvrC from E.coli.* EMBO J, 2002. **21**(22): p. 6257-66.
38. Easton, A.M. and S.R. Kushner, *Transcription of the uvrD gene of Escherichia coli is controlled by the lexA repressor and by attenuation.* Nucleic Acids Res, 1983. **11**(24): p. 8625-40.
39. Maples, V.F. and S.R. Kushner, *DNA repair in Escherichia coli: identification of the uvrD gene product.* Proc Natl Acad Sci U S A, 1982. **79**(18): p. 5616-20.
40. Orren, D.K., et al., *Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I.* J Biol Chem, 1992. **267**(2): p. 780-8.
41. Gross, J. and M. Gross, *Genetic analysis of an E. coli strain with a mutation affecting DNA polymerase.* Nature, 1969. **224**(5225): p. 1166-8.
42. Campbell, J.L., L. Soll, and C.C. Richardson, *Isolation and partial characterization of a mutant of Escherichia coli deficient in DNA polymerase II.* Proc Natl Acad Sci U S A, 1972. **69**(8): p. 2090-4.
43. Scharer, O.D., *Nucleotide excision repair in eukaryotes.* Cold Spring Harb Perspect Biol, 2013. **5**(10): p. a012609.

44. Song, J., M.G. Kemp, and J.H. Choi, *Detection of the Excised, Damage-containing Oligonucleotide Products of Nucleotide Excision Repair in Human Cells.* Photochem Photobiol, 2017. **93**(1): p. 192-198.
45. Lindahl, T., *An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues.* Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3649-53.
46. Kim, Y.J. and D.M. Wilson, 3rd, *Overview of base excision repair biochemistry.* Curr Mol Pharmacol, 2012. **5**(1): p. 3-13.
47. Krokan, H.E., et al., *Base excision repair of DNA in mammalian cells.* FEBS Lett, 2000. **476**(1-2): p. 73-7.
48. Seeberg, E., L. Eide, and M. Bjoras, *The base excision repair pathway.* Trends Biochem Sci, 1995. **20**(10): p. 391-7.
49. Friedberg, E.C., et al., *DNA repair and mutagenesis.* 2006: p. 169-226.
50. Alexeeva, M., et al., *Excision of uracil from DNA by hSMUG1 includes strand incision and processing.* Nucleic Acids Res, 2019. **47**(2): p. 779-793.
51. Whitaker, A.M., et al., *Base excision repair of oxidative DNA damage: from mechanism to disease.* Front Biosci (Landmark Ed), 2017. **22**: p. 1493-1522.
52. Kavli, B., et al., *hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup.* J Biol Chem, 2002. **277**(42): p. 39926-36.
53. Nilsen, H., et al., *Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase.* EMBO J, 2001. **20**(15): p. 4278-86.
54. Pearl, L.H., *Structure and function in the uracil-DNA glycosylase superfamily.* Mutat Res, 2000. **460**(3-4): p. 165-81.
55. Bjelland, S., M. Bjoras, and E. Seeberg, *Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of Escherichia coli.* Nucleic Acids Res, 1993. **21**(9): p. 2045-9.
56. Doetsch, P.W. and R.P. Cunningham, *The enzymology of apurinic/apyrimidinic endonucleases.* Mutat Res, 1990. **236**(2-3): p. 173-201.
57. Demple, B., A. Johnson, and D. Fung, *Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H2O2-damaged Escherichia coli.* Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7731-5.
58. Cunningham, R.P., et al., *Endonuclease IV (nfo) mutant of Escherichia coli.* J Bacteriol, 1986. **168**(3): p. 1120-7.
59. Dianov, G. and T. Lindahl, *Reconstitution of the DNA base excision-repair pathway.* Curr Biol, 1994. **4**(12): p. 1069-76.
60. Xu, G., et al., *Base excision repair, aging and health span.* Mech Ageing Dev, 2008. **129**(7-8): p. 366-82.
61. Cupples, C.G. and J.H. Miller, *A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions.* Proc Natl Acad Sci U S A, 1989. **86**(14): p. 5345-9.
62. Garibyan, L., et al., *Use of the rpoB gene to determine the specificity of base substitutions on the Escherichia coli chromosome.* DNA Repair, 2003: p. 593-608.
63. Rosche, W.A. and P.L. Foster, *Determining mutation rates in bacterial populations.* Methods, 2000. **20**(1): p. 4-17.
64. Luria, S.E. and M. Delbruck, *Mutations of Bacteria from Virus Sensitivity to Virus Resistance.* Genetics, 1943. **28**(6): p. 491-511.
65. Bjelland, S., et al., *DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in Escherichia coli.* J Biol Chem, 1994. **269**(48): p. 30489-95.
66. Lin, J.J. and A. Sancar, *(A)BC excinuclease: the Escherichia coli nucleotide excision repair enzyme.* Mol Microbiol, 1992. **6**(16): p. 2219-24.
67. Sancar, A., *DNA excision repair.* Annu Rev Biochem, 1996. **65**: p. 43-81.

68. Van Houten, B. and A. Snowden, *Mechanism of action of the Escherichia coli UvrABC nuclease: clues to the damage recognition problem.* Bioessays, 1993. **15**(1): p. 51-9.

Appendices

A.1 Effects of the UvrABC system on mutation induction by 5-formyldeoxyuridine in Escherichia coli.

A.1.1 Buffers, solutions and dishes

Table A.1.1: 10× A buffer

Components	Amount used in preparation	Manufacturer
K_2HPO_4	105 g	Merck: 1.05104.1000, Mw: 174.18 g/mol
KH_2PO_4	45 g	Merck: 1.04873.1000, Mw: 136.09 g/mol
$(NH_4)_2SO_4$	10 g	Merck: 1.01217.1000, Mw: 132.10 g/mol
$C_6H_5Na_3O_7$ x 2 H_2O	5g	Merck: 1.06448.1000, Mw: 294.10 g/mol
dH_2O	Dilute to 1000 ml	

Autoclave before use

Table A.1.2: 50× TAE

Components	Amount used in preparation	Manufacturer
Tris Base	48.4 g	Sigma, Cat. #T6066, 121.14 g/mol
99.5% Acetic acid	11.42 ml	
0.5 M EDTA, pH 8	20 ml	Lab stock
Deionized H_2O	Dilute to 200 ml	

Autoclave before use

Table A.1.3: 1 M $MgSO_4$

Components	Amount used in preparation	Manufacturer
$MgSO_4 \times 7H_2O$	22.85 g	Merck: 1.05886.0500, Mw: 246.48 g/mol
Deionized H_2O	Dilute to 100 ml	

Autoclave before use, store at 4°C in a dark bottle.

Table A.1.4: 0.5 M EDTA, pH 8

Components	Amount used in preparation	Manufacturer
EDTA	37.22 g	Sigma: ED2SS, Mw: 372.2 g/mol
Deionized H_2O	Dilute to 200 ml*	

Dilute to 150 ml, and then adjust pH to dissolve the EDTA, using hydrochloric acid.
Autoclave before use.

Table A.1.5: 20% Glucose

Components	Amount used in preparation	Manufacturer
Glucose	40.0 g	Merck: 1.08337.1000, Mw: 180.16 g
Deionized H_2O	Dilute to 200 ml	

Autoclave before use, store at 4°C.

Table A.1.6: Vitamine B_1 (5 mg/ml)

Components	Amount used in preparation	Manufacturer
Thiamine	0.05 g	Sigma: T4625-10G, Mw 337.27 g/mol
Deionized H_2O	Dilute to 10 ml	

Autoclave before use.

Table A.1.7: L-amino acids (4 mg/ml)

Components	Amount used in preparation	Manufacturer
L-threonine	1 g	Sigma: T-8441, Mw: 119.1 g/mol
L-arginine	1 g	Sigma: A-5131, Mw: 210.7 g/mol
L-proline	1 g	Sigma: P-0380, Mw: 115.1 g/mol
L-leucine	1 g	Sigma: L-8125, Mw: 209.6 g/mol
L-histidine	1 g	Sigma: H-8125, Mw: 155.15 g/mol
Deionized H_2O	Dilute to 250 ml	

Autoclave before use.

Table A.1.8: Rifampicin (30 mg/ml)

Components	Amount used in preparation	Manufacturer
Rifampicin	0.150 g	Sigma: R-3501, Mw: 823 g/mol
Methanol	Dilute to 5 ml	

Store at -20°C and add to LB media when cultivating mutant colonies. Make fresh solution when adding to media to make dishes.

Table A.1.9: f^5dU (10 mM)

Components	Amount used in preparation	Manufacturer
f^5dU	0.026 g	Gift from prof A.Matsuda, Mw: 256.18 g/mol
Deionized H_2O	Dilute to 10 ml	

Store in aliquotes at -20°C.

Table A.1.10: A-medium (liquid)

Components	Amount used in preparation
10× A buffer	20 ml
1 M $MgSO_4$	0.2 ml
20% Glucose	2 ml
Amino acids (4 mg/ml)	2 ml
Vitamin B_1 (5 mg/ml)	0.2 ml
Deionized H_2O	Dilute to 200 ml

Autoclave before use, store at 4°C.

Table A.1.11: Glucose dishes

Components	Amount used in preparation
Agar	12 g
10× A buffer	100 ml
20% Glucose	10 ml
1 M $MgSO_4$	1 ml
Amino acids (4 mg/ml)	1 ml
Vitamin B_1 (5 mg/ml)	1 ml
Deionized H_2O	Dilute to 1000 ml

Autoclave and transfer to dishes.

To make glucose plates with rifampicin, let the bottle cool down to 55°C and add 5 ml rifampicin solution (30 mg/ml) before transferring to dishes. Store for up to one month at 4°C.

Table A.1.12: Agarose gel

Components	Amount used in preparation	Manufacturer
Agarose	0.5 g	Sigma: 5093
1× TAE	Dilute to 50 ml	

Boil in microwave until the agarose is completely dissolved. Cool down to 55°C before adding 5 µl of GelRed. Mix well and pour into chamber. Let dry for 20-30 min prior to use.

A.1.2 Detailed protocols

A.1.2.1 Mutagenesis

1 – Starting overnight (ON) culture in liquid medium
The bacterial cultures are stored at -80°C. Start an ON-culture by using a sterile loop (1 µl) to transfer bacteria to a sterile falcon tube containing LB medium (5 ml), followed by incubation for 18-24 h (37°C, 240 rpm).

2 – Spreading bacteria from overnight culture
Use a sterile plastic loop (1 µl) to spread the bacteria onto a LB plate, which functiones as a stock where colonies for further analysis are harvested. Incubate the plate for 18-24 h at 37°C, and then store at 4°C. To keep the colonies fresh, a single colony should be harvested and spread onto a new LB plate approximately once a week.

3 – Starting ON cultures for analysis
Cultures are harvested from the stock plate and transferred into liquid A-medium (2 ml). For each assay six colonies are transferred to six different culture tubes, before incubation at 37°C for 18-24 hours at a tilted angle. One control tube containing only liquid A-medium should also be incubated under the same conditions.

4 – Measuring concentration and diluting cultures
Use the Bio-Rad spectrophotometer to measure absorption of the cultures at 600 nm. The control sample without added colony is used to blank the spectrophotometer. The ideal absorption values are between 0.4 and 0.8, so the samples with a value closest to 0.6 should be chosen for further analysis.
The cultures are then diluted to 200.000 bact/ml by first diluting 1:100 by adding 10 µl of culture to 990 µl 1×A-buffer, and then calculate the amount of 1×A-buffer needed for further dilution using this formula: (bact/ml ($\times 10^8$)×10×25)-25=amount of 1×A-buffer needed. The 1:100 dilution (50 ml) is then added to the calculated amount of A-buffer, and 293 µl of the final dilution is added to a falcon tube containing A-medium (6,5 ml), and incubated for 2h at 37°C and 240 rpm to adapt cells, and ensure that mutagenesis is performed during exponential growth.

5 – Mutagenesis
After two hours each sample is divided into three parallels (2 mL), and f^5dU is added (20 µl, sterile) to each parallel, for control samples no f^5dU is added. Further incubation followes at 37°C, 240 rpm for 45-48h.

6 – Harvesting cells
Put the cultures on ice to stop the mutagenesis, and vortex each glass tube before dividing into two Eppendorf tubes (1,5 ml microspin tube, 1 ml each). The tubes are then centrifuged; 5000 rpm, 4

min, 4°C. The supernatant is discarded, and 1× A-buffer (1 ml) is used to wash the pellet. Perform a second centrifugation with the same settings and discard the supernatant again. Resuspend the pellets in 1× A-buffer (0,5 ml to each tube) before the two tubes of each samples are combined again (giving a total of 1 tube with 1 ml washed bacteria per sample).

7 – Dilution of bacterial cultures
The bacterial cultures are diluted as described in table A.1.13:

Table A.1.13: Dilution scheme for bacterial cultures

Dilution		Bacterial culture	1× A-buffer	Total volume
10^{-2}	1:100	10 µl concentrated	990 µl	1 000 µl
10^{-4}	1:10 000	10 µl 1:100 dilution	990 µl	1 000 µl
10^{-5}	1:100 000	100 µl 1:10 000 dilution	900 µl	1 000 µl
10^{-6}	1:1000 000	100 µl 1:100 000 dilution	900 µl	1 000 µl
10^{-7}	1:10 000 000	100 µl 1:1 000 000 dilution	900 µl	1 000 µl

8 – Spreading dishes
For each assay, a total of 60 minimal media (glucose) plates, and 72 minimal media + rifampicin plates are used. The dishes should be marked with assay number, sample number, dilution and initials prior to use. Add cultures and spread bacteria using an ethanol (70%) and flame sterilized steel loop according to table A.1.14. Incubate the dishes at 37°C for 48 hours (minimal media) or 96 hours (minimal media + rifampicin).

Table A.1.14: Plates used for mutagenesis assay

Dilution	Volume	Dish
10^{-5}	100 µl	Minimal media ×1
10^{-6}	100 µl	Minimal media ×2
10^{-7}	100 µl	Minimal media ×2
Concentrated culture	100 µl	Minimal media + rifampicin ×6

9 – Cultivating mutant colonies
For every sample, a suitable colony should be harvested and inoculated into LB media (2 ml), containing 150 µg/ml rifampicin, and vortexed before incubating at 37°C, 240 rpm for 5-7days.

A.1.2.2 DNA-extraction, -amplification and -sequencing

10 – DNA extraction
Put the cultures on ice, before mixing with sterile water (5 µl culture and 100µl water, DNase-, RNase-, and -Protease free). Boil the mixture for 5 min, before cooling on ice (1 min) prior to centrifugation; 13 000 rpm, 3 min, 4°C. The supernatant (80 µl) is then transferred to a sterile microspin tube (1,5 ml).

11 – Polymerase Chain Reaction (PCR)
The PCR reaction is set up as described in table XXX.

Table A.1.15: PCR reaction mix

Reagent	1x Reaction (µl)	Final concentration
5× GoTaq Flexi buffer	10	1×
25 mM $MgCl_2$	3	1.5 mM
dNTP	1	200 µM each
Primer 1021 (10 pmol/µl)	1	0.2 pmol/µl
Primer 1022 (10 pmol/µl)	1	0.2 pmol/µl
GoTaq Hot start DNA polymerase (5 U/µl)	0.25	1.25 U/µl
Sterile H_2O	28.75	
Total	45	

For each PCR, multiply the reaction mix with the number of samples (n) +2 (one negative control, and one extra).
Aliquot he reaction mix (45 µl) into n+1 microspin tubes (0,2 ml) and add the DNA templates (5 µl). Sterile water is used as a negative control.
Amplify the DNA using the parameters in table A.1.16.

Table A.1.16: PCR set up

Step	Time	Temperature	Cycles
Initial denaturation	4 min	94°C	×1
Denaturation	1 min	94°C	×4
Annealing	1 min	50°C	
Extension	20 sec	72°C	
Final extension	5 min	72°C	×1
Storage	∞	4°C	

12 – Gel of PCR products
Run the PCR products on a 1% agarose gel with GelRed (5 µl) for 40min at 100V

13 – Purification of PCR products
When necessary the PCR products should be purified using a purification kit (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel).

1. 2×V NT1 is added to the PCR product (45 µl PCR products=90 µl NT1), mixed and then 90 µl should be transferred to the column, which is placed in a collecting tube.
2. The columns are centrifuged; 11 000 rpm, 30 sec, 4°C.
3. Discard the content in the collection tube.
4. NT3 buffer is added to the column (700 µl).
5. The columns are centrifuged; 11 000 rpm, 30 sec, 4°C.

6. The content in the collection tube is discarded.
7. Steps 4 through 6 are repeated.
8. Dry the silica in the column; 11 000 rpm, 30 sec, 4°C.
9. Discard the content in the collection tube.
10. Place the column in a regular microspin tube (1,5 ml).
11. Elute the DNA by adding NE (15 µl), and incubate for 1 minute.
12. Centrifuge the columns; 1 .000 rpm, 1 min, 4°C.

14 – Nanodrop measure
The surface of the instrument should first be cleaned using sterile water (DNase-, RNase-, and -Protease free), before calibrating with either sterile water or NE (sterile water for PCR products, NE for purified PCR samples). Samples are measured (2 µl), and results registered.

15 – Preparation for sequencing
For each sample, dilute 5 µl PRC products (or purified PCR products) with primer mix (5 µl), before shipping to sequencing. The primer mix consists of 5 µl 1021 primer 10 mM and 5µl sterile water (DNase-, RNase-, and -Protease free).
Lable the tubes with barcodes, and insert the corresponding sticker to the right sample in your journal.
Sequencing is performed by GATC Biotech AG, European custom sequencing centre, using sanger sequencing.

A.1.3 Complete assay overview

Assay number: 2019/1

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 19 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.394	1.97	d	0.502	2.51
b	0.010	0	e	0.459	2.29
c	0.504	2.52	f	0.512	2.56

Cultures chosen for further analysis: c=A, d=B, e=C, f=D

Mutagenesis: 45.5 h

Notes

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	-	-	-	-	-
AF I	-	-	-	-	-
AF II	-	-	-	-	-
BK	-	-	-	-	-
BF I	-	-	-	-	-
BF II	-	-	-	-	-
CK	-	-	-	-	-
CF I	1	-	-	-	-
CF II	-	-	-	-	-
DK	2	2	1	-	-
DF I	-	-	-	-	-
DF II	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	-	-	6
AF II	-	-	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	-	-	-	-	-	-	6
BF II	-	-	-	-	-	-	6
CK	-	-	-	-	-	-	6
CF I	-	-	-	-	-	-	6
CF II	-	-	-	-	-	-	6
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II	-	-	-	-	-	-	6

Notes: Fire in lab during critical step of the assay. Work was interrupted and is probably the reason for no growth in the dishes.
Incubation time glucose dishes: 50.5 h Incubation time rifampicin dishes: 91 h
No mutant colonies to cultivate, assay aborted at this stage.

Assay number: 2019/2

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 19.75 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.513	2.56	d	0.513	2.57
b	0.522	2.61	e	0.525	2.63
c	0.541	2.71	f	0.527	2.63

Cultures chosen for further analysis: b=A, c=B, e=C, f=D

Mutagenesis: 46 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	9	1	-	-	-
AF I	-	-	-	-	-
AF II	-	-	-	-	-
BK	-	-	-	-	-
BF I	260	16	29	4	-
BF II	136	17	7	1	3
CK	-	-	-	-	-
CF I	217	26	15	2	2
CF II	-	-	-	-	-
DK	-	-	-	-	-
DF I	-	-	-	-	-
DF II	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK	**1**	**-**	**-**	**-**	**-**	**-**	6
AF I	**-**	**-**	**-**	**-**	**-**	**-**	6
AF II	**-**	**-**	**-**	**-**	**-**	**-**	6
BK	**-**	**-**	**-**	**-**	**-**	**-**	6
BF I	**-**	**-**	**-**	**-**	**-**	**-**	6
BF II	**1**	**-**	**-**	**-**	**-**	**-**	6
CK	**-**	**-**	**-**	**-**	**-**	**-**	6
CF I	**1**	**-**	**-**	**-**	**-**	**-**	6
CF II	**-**	**-**	**-**	**-**	**-**	**-**	6
DK	**-**	**-**	**-**	**-**	**-**	**-**	6
DF I	**-**	**-**	**-**	**-**	**-**	**-**	6
DF II	**-**	**-**	**-**	**-**	**-**	**-**	6

Notes: Incubation time glucose dishes: 49 h Incubation time rifampicin dishes: 96 h
3 colonies from three different plates were cultivated in LB medium containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/3

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21,5 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.578	2.89	d	0.476	2.38
b	0.551	2.76	e	0.541	2.71
c	0.578	2.89	f	0.533	2.66

Cultures chosen for further analysis: a=A, b=B, c=C, e=D

Mutagenesis: 45 h

Notes: No growth during mutagenesis in tubes CF II, DF I and DF II. These samples were not included in further analysis. Contamination in one of the Eppendorf tubes in test CK, only half volume in sample, therefore only 4 rifampicin dishes.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	14	-	2	-	-
AF I	4	-	-	-	-
AF II	11	3	-	-	-
BK	2	1	-	-	-
BF I	89	7	-	-	-
BF II	30	1	-	-	-
CK	2	1	2	-	-
CF I	1	1	-	-	-
CF II					
DK	5	3	1		1
DF I					
DF II					

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	-	-	6
AF II	-	-	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	-	-	**2**	**3**	-	**95**	6
BF II	-	-	-	-	-	-	6
CK	-	-	-	-			4
CF I	-	-	-	-	-	-	6
CF II							
DK	-	-	-	-	-	-	6
DF I							
DF II							

Notes: Incubation time glucose dishes: 48.5 h Incubation time rifampicin dishes: 91.5 h
1 colony from 1 sample was cultivated in LB medium containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/4

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 22 hours

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.491	2.45	d	0.436	2.18
b	0.451	2.26	e	0.459	2.30
c	0.412	2.06	f	0.468	2.34

Cultures chosen for further analysis: a=A, b=B, e=C, f=D

Mutagenesis: 45.5 h

Notes: Overnight cultures grown in plastic tubes, no glass tubes available.
No growth during mutagenesis for cultures AF I, AF II, BF I, BF II, CF I and CF II. These samples were not included in further analysis.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	21	8	3	-	-
AF I					
AF II					
BK	1	2	7	-	-
BF I					
BF II					
CK	7	1	-	-	-
CF I					
CF II					
DK	28	5	-	-	-
DF I	82	10	1	1	1
DF II	68	3	1	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	**4**	-	6
AF I							
AF II							
BK	-	-	-	-	-	-	6
BF I							
BF II							
CK	-	-	-	-	**1**	-	6
CF I							
CF II							
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II	-	-	**1**	-	-	-	6

Notes: Incubation time glucose dishes: 49.5 h Incubation time rifampicin dishes: 96 h
3 colonies from 3 different samples cultivated in LB medium containing 30 mg/ml rifampicin for 6 days.

Assay number: 2019/5

Bacterial strain: AB 1885 ($uvrB^-$)

Overnight culture: 23,5 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.524	2.62	d	0.454	2.26
b	0.532	2.66	e	0.526	2.63
c	0.527	2.64	f	0.521	2.60

Cultures chosen for further analysis: a=A, b=B, c=C, e=D

Mutagenesis: 45 h

Notes: No growth in any tubes during mutagenesis, assay aborted at this point.

Assay number: 2019/6

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 20 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.452	2.26	e	0.631	3.15
b	0.536	2.68	f	0.591	2.96
c	0.416	2.08	g	0.622	3.11
d	0.571	2.86			

Cultures chosen for further analysis: d=A, e=B, f=C, g=D

Mutagenesis: 45 h

Notes: No growth in sample AF I, sample excluded from further analysis.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	50	62	66	13	29
AF I					
AF II	61	6	3	-	1
BK	509	89	89	12	26
BF I	259	35	52	6	7
BF II	516	89	93	14	1
CK	546	59	55	6	8
CF I	457	44	55	7	9
CF II	411	57	62	7	7
DK	531	82	67	8	6
DF I	370	54	53	7	6
DF II	414	56	49	4	4

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I							
AF II	-	-	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	**150**	**172**	**164**	**188**	**146**	**145**	6
BF II	**1**	-	**2**	**1**	-	**1**	6
CK	-	**1**	-	-	-	-	6
CF I	-	**1**	**1**	**1**	-	-	6
CF II	-	-	-	-	**2**	**2**	6
DK	**3**	**2**	**2**	**3**	**4**	**1**	6
DF I	-	-	**3**	**1**	**1**	**1**	6
DF II	-	**1**	**3**	**1**	**1**	**1**	6

Notes: Incubation time glucose dishes: 48 h Incubation time rifampicin dishes: 92 h
8 colonies from 8 different samples were cultivated in LB medium containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/7

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.628	3.14	d	0.631	3.16
b	0.627	3.14	e	0.638	3.19
c	0.629	3.14	f	0.622	3.11

Cultures chosen for further analysis: a=A, b=B, c=C, f=D

Mutagenesis: 45.5 h

Notes: No visible turbidity in sample DF II, no pellet after centrifugation. Sample excluded from the rest of the assay.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	2	-	1	-	-
AF I	-	-	-	-	-
AF II	3	-	-	-	-
BK	-	-	-	-	-
BF I	4	2	3	-	-
BF II	-	-	-	-	-
CK	-	-	-	-	-
CF I	-	-	-	-	-
CF II	-	-	-	-	-
DK	-	-	-	-	-
DF I	1	-	-	-	-
DF II					

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	-	-	6
AF II	-	-	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	-	-	-	-	-	-	6
BF II	-	-	-	-	-	-	6
CK	-	-	-	-	-	-	6
CF I	-	-	-	-	-	-	6
CF II	-	-	-	-	-	-	6
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II							

Notes: Incubation time glucose dishes: 50.5 h Incubation time rifampicin dishes: 93.5 h
No mutant colonies, assay aborted at this stage.

Assay number: 2019/8

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 22.5 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.661	3.30	d	0.647	3.23
b	0.638	3.19	e	0.661	3.31
c	0.631	3.15	f	0.646	3.23

Cultures chosen for further analysis: b=A, c=B, d=C, f=D

Mutagenesis: 45.5 h

Notes: No visible growth, and no pellet formed during centrifugation in the following samples: AF II, BK, BF II, DF I and DF II. These samples were excluded from the rest of the assay.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	602	103	86	8	10
AF I	517	48	78	9	5
AF II					
BK					
BF I	635	81	78	7	12
BF II					
CK	746	93	112	11	10
CF I	586	56	74	3	4
CF II	380	75	61	6	8
DK	413	59	26	8	11
DF I					
DF II					

Number of colonies on rifampicin dishes							**Number of dishes**
AK	**-**	**-**	**-**	**-**	**-**	**-**	6
AF I	**2**	**1**	**-**	**1**	**-**	**-**	6
AF II							
BK							
BF I	**3**	**4**	**3**	**4**	**6**	**5**	6
BF II							
CK	**8**	**7**	**14**	**6**	**10**	**14**	6
CF I	**1**	**-**	**2**	**3**	**-**	**2**	6
CF II	**-**	**1**	**1**	**-**	**1**	**-**	6
DK	**4**	**3**	**6**	**1**	**1**	**1**	6
DF I							
DF II							

Notes: Incubation time glucose dishes: 51 h Incubation time rifampicin dishes: 9.,5 h
6 colonies from 6 different samples were cultivated in LB medium containing 30 mg/ml rifampicin for 6 days

Assay number: 2019/9

Bacterial strain: AB 1885 (*uvrB*$^{-}$)

Overnight culture: 21 hours

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.617	3.08	d	0.631	3.15
b	0.655	3.27	e	0.653	3.26
c	0.629	3.14	f	0.638	3.19

Cultures chosen for further analysis: a=A, c=B, d=C, f=D
Mutagenesis:
Notes: Assay aborted during mutagenesis due to illness.

Assay number: 2019/10

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.661	3.30	d	0.686	3.43
b	0.645	3.23	e	0.659	3.29
c	0.659	3.30	f	0.679	3.40

Cultures chosen for further analysis: a=A, b=B, c=C, e=D

Mutagenesis: 47.5 h

Notes: Due to unresolved growth issues during previous assays, growth period prior to mutagenesis was extended from 2 to 3 h in this assay. This is the first assay with visible turbidity and pellet in all samples.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	609	72	96	5	12
AF I	657	88	69	8	9
AF II	526	67	50	6	7
BK	603	77	59	5	8
BF I	500	3	73	-	-
BF II	501	78	52	17	4
CK	526	63	42	5	2
CF I	466	77	83	7	8
CF II	494	71	70	8	4
DK	-	-	-	-	-
DF I	347	29	65	6	2
DF II	411	38	55	4	7

Number of colonies on rifampicin dishes							**Number of dishes**
AK	**-**	**-**	**-**	**1**	**-**	**-**	6
AF I	**-**	**1**	**1**	**1**	**2**	**-**	6
AF II	**1**	**2**	**3**	**12**	**-**	**1**	6
BK	**1**	**-**	**1**	**-**	**2**	**-**	6
BF I	**4**	**6**	**-**	**8**	**2**	**5**	6
BF II	**-**	**3**	**-**	**2**	**1**	**1**	6
CK	**-**	**-**	**39**	**-**	**2**	**2**	6
CF I	**-**	**-**	**-**	**-**	**-**	**2**	6
CF II	**1**	**-**	**-**	**-**	**-**	**-**	6
DK	**-**	**-**	**-**	**-**	**-**	**-**	6
DF I	**8**	**2**	**4**	**2**	**4**	**1**	6
DF II	**1**	**-**	**-**	**-**	**-**	**1**	6

Notes: Incubation time glucose dishes: 48 h Incubation time rifampicin dishes: 94 h
11 colonies from 11 different samples were cultivated in LB medium containing 30 mg/ml rifampicin for 7 days

Assay number: 2019/11

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.616	3.08	d	0.577	2.88
b	0.578	2.89	e	0.578	2.89
c	0.617	3.08	f	0.563	2.82

Cultures chosen for further analysis: a=A, b=B, c=C, e=D

Mutagenesis: 45 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	480	47	65	2	5
AF I	522	46	39	10	9
AF II	605	58	50	9	6
BK	469	51	53	10	6
BF I	500	54	54	6	9
BF II	373	48	37	8	1
CK	461	45	36	2	5
CF I	348	41	41	5	4
CF II	346	49	29	2	6
DK	324	43	35	4	3
DF I	257	17	10	1	2
DF II	339	34	33	2	3

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	**2**	-	6
AF II	-	-	-	-	**3**	**4**	6
BK	**1**	-	-	**1**	**1**	**1**	6
BF I	**1**	**1**	-	-	-	-	6
BF II	-	-	-	-	-	-	6
CK	-	-	-	-	-	-	6
CF I	-	-	-	-	-	-	6
CF II	-	-	-	-	-	-	6
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II	-	-	**1**	**1**	-	-	6

Notes: Incubation time glucose dishes: 50 h Incubation time rifampicin dishes: 95
5 colonies from 5 different samples were cultivated in LB medium containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/12

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.528	2.64	d	0.534	2.67
b	0.540	2.70	e	0.568	2.84
c	0.512	2.56	f	0.546	2.73

Cultures chosen for further analysis: b=A, d=B, e=C, f=D

Mutagenesis: 46 h

Notes: The settings on the shaking incubator was changed sometime during the incubation period from 240 rpm to 200 rpm. I do not think if that this is relevant to the survival rate of the bacteria. There was some turbidity in the tubes, and small visible pellets after centrifugation, but almost no growth. I am not sure if any other settings were changed during the mutagenesis, as this is a shared resource.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	2	-	-	-	-
AF I	2	-	-	-	-
AF II	1	-	-	-	-
BK	-	-	-	-	-
BF I	-	-	-	-	-
BF II	-	-	-	-	-
CK	1	-	-	-	1
CF I	-	-	-	-	-
CF II	-	-	-	-	-
DK	-	-	-	-	-
DF I	-	-	-	-	-
DF II	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	-	-	6
AF II	-	-	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	-	-	-	-	-	-	6
BF II	-	-	-	-	-	-	6
CK	-	-	-	-	-	-	6
CF I	-	-	-	-	-	-	6
CF II	-	-	-	-	-	-	6
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 48 h Incubation time rifampicin dishes: 95 h
No mutant colonies, assay aborted at this stage.

Assay number: 2019/13

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21.5 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.580	2.90	d	0.563	2.81
b	0.576	2.88	e	0.573	2.87
c	0.522	2.76	f	0.561	2.80

Cultures chosen for further analysis: a=A, b=B, d=D, e=D

Mutagenesis: 45 h

Notes: The settings on the shaking incubator was changed sometime during the incubation period from 240 rpm to 200 rpm. I do not think that this is relevant to the survival rate of the bacteria. There was some turbidity in the tubes, and small visible pellets after centrifugation, but almost no growth. I am not sure if any other settings were changed during the mutagenesis, as this is a shared resource.

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	-	-	-	-	-
AK II	-	-	-	-	-
AK III	1	-	-	-	-
BK I	2	-	-	-	-
BK II	-	-	-	-	-
BK III	2	-	-	-	-
CK I	-	1	1	-	-
CK II	-	-	-	-	-
CK III	-	-	-	-	-
DK I	-	-	1	-	-
DK II	-	-	-	-	-
DK III	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 48 h Incubation time rifampicin dishes: 94 h
No mutant colonies, assay aborted at this stage.

Assay number: 2019/14

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.570	2.85	d	0.582	2.91
b	0.584	2.92	e	0.593	2.97
c	0.612	3.06	f	0.577	2.89

Cultures chosen for further analysis: b=A, c=B, d=C, e=D

Mutagenesis: 45 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	-	-	-	-	-
AK II	-	-	-	-	-
AK III	1	-	-	-	-
BK	2	-	-	-	-
BK II	-	-	-	-	-
BK III	2	-	-	-	-
CK I	-	1	1	-	-
CK II	-	-	-	-	-
CK III	-	-	-	-	-
DK I	-	-	1	-	-
DK II	-	-	-	-	-
DK III	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 50 h Incubation time rifampicin dishes: 93 h
No mutant colonies, assay aborted at this stage.

Assay number: 2019/15

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.570	2.85	d	0.582	2.91
b	0.584	2.92	e	0.593	2.97
c	0.612	3.06	f	0.577	2.89

Cultures chosen for further analysis: b=A, c=B, d=C, e=D

Mutagenesis: 45 h

Notes: To check whether media was involved in the major growth challenges experienced, LB was added to some plates in this assay, as indicated by * below.

Number of colonies on glucose dishes								
Dish	10^{-5}		10^{-6}			10^{-7}		
AK I	41	716*	15	-	69*	-	-	9*
AK II	1	816*	5	9	112*	-	-	11*
AK III	36	638*	-	1	68*	-	-	10*
BK I	180	552*	-	2	102*	-	-	14*
BK II	397	532*	1	17	73*	-	-	11*
BK III	128	549*	-	4	94*	-	-	3*
CK I	-		-	-		-	-	
CK II	-		-	-		-	-	
CK III	-		-	-		-	-	
DK I	-		1	-		-	-	
DK II	1		-	-		-	-	
DK III	1		-	-		-	-	

*LB dishes

Number of colonies on rifampicin dishes									**Number of dishes**
AK I	**1**	**-**	**-**	**-**	**-**	**-**	**-***	**2***	8
AK II	**-**	**-**	**-**	**-**	**-**	**-**	**-***	**-***	8
AK III	**-**	**-**	**-**	**-**	**-**	**-**	**16***	**18***	8
BK I	**-**	**-**	**-**	**-**	**-**	**-**	**7***	**4***	8
BK II	**-**	**-**	**-**	**-**	**-**	**-**	**8***	**8***	8
BK III	**-**	**-**	**-**	**-**	**-**	**-**	**-***	**-***	8
CK I	-	-	-	-	-	-			6
CK II	-	-	-	-	-	-			6
CK III	-	-	-	-	-	-			6
DK I	-	-	-	-	-	-			6
DK II	-	-	-	-	-	-			6
DK III	-	-	-	-	-	-			6

*LB+rifampicin dishes

Notes: Incubation time glucose dishes: 73.5 h
LB dishes incubated for 24 h
Incubation time rifampicin dishes: 116.5 h.
Incubation time LB+rifampicine dishes: 74 h.
5 colonies from 5 different dishes were cultivated in LB+rifampicine media for 6 days.

Assay number: 2019/16

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.510	2.55	d	0.511	2.56
b	0.502	2.51	e	0.556	2.78
c	0.513	2.57	f	0.554	2.77

Cultures chosen for further analysis: c=A, d=B, e=C, f=D

Mutagenesis: 45h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	-	-	-	-	-
AK II	-	-	-	-	-
AK III	-	-	-	-	-
BK I	-	-	-	-	-
BK II	-	-	-	-	-
BK III	-	-	-	-	-
CK I	-	-	-	-	-
CK II	-	-	-	-	-
CK III	-	-	-	-	1
DK I	-	-	-	-	-
DK II	-	-	-	-	-
DK III	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 93.5 h Incubation time rifampicin dishes: 96 h
No mutant colonies, assay aborted at this stage

Assay number: 2019/17

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.529	2.65	d	0.581	2.90
b	0.589	2.95	e	0.624	3.12
c	0.613	3.07	f	0.599	3.00

Cultures chosen for further analysis: b=A, c=B, d=C, f=D

Mutagenesis: 45 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	-	-	1	-	-
AK II	-	-	-	-	-
AK III	-	-	-	-	-
BK I	-	-	-	-	-
BK II	-	-	-	-	-
BK III	1	-	1	-	-
CK I	-	-	-	-	-
CK II	-	-	-	-	-
CK III	-	-	-	-	-
DK I	-	-	-	-	-
DK II	-	-	-	-	-
DK III	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 92.5 h Incubation time rifampicin dishes: 92.5 h
No mutant colonies, assay aborted at this stage.

Assay number: 2019/18

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.630	3.15	d	0.659	3.30
b	0.618	3.09	e	0.563	2.82
c	0.629	3.14	f	0.603	3.02

Cultures chosen for further analysis: a=A, b=B, c=C, f=D

Mutagenesis: 45 h

Notes:

Number of colonies on glucose dishes						
Dish	10^{-5}	10^{-6}		10^{-7}		
AK I	28	9	1	-	-	
AK II	19	6	-	-	1	
AK III	11	-	2	1	1	
BK I	26	18	22	4	2	
BK II	27	15	23	1	1	
BK III	14	2	5	-	-	
CK I	16	2	2	-	1	
CK II	12	-	-	1	-	
CK III	11	1		3	1	1
DK I	4	-	-	-	1	
DK II	27	3	1	1	1	
DK III	4	-	-	-	-	

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	**2**	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	**1**	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 69 h Incubation time rifampicin dishes: 95 h
2 colonies from 2 different samples were cultivated in LB medium containing 30 mg/ml rifampicin for 7 days

Assay number: 2019/19

Bacterial strain: AB 1884 (*uvrC*$^-$)

Overnight culture: 21.5 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)	Culture nr	OD at 600 nm	Bact/ml ($\times 10^8$)
a	0.629	3.15	d	0.644	3.22
b	0.630	3.15	e	0.650	3.25
c	0.632	3.16	f	0.631	3.15

Cultures chosen for further analysis: a=A, b=B, c=C, f=D

Mutagenesis: 45 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK	57	6	14	1	-
AF I	10	3	6	-	-
AF II	11	4	4	-	-
BK	7	3	3	-	-
BF I	22	16	10	-	3
BF II	3	-	2	-	-
CK	6	1	15	-	-
CF I	10	1	5	-	2
CF II	25	6	7	-	1
DK	2	8	3	1	-
DF I	64	10	3	-	3
DF II	6	-	1	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK	-	-	-	-	-	-	6
AF I	-	-	-	-	-	-	6
AF II	-	**1**	-	-	-	-	6
BK	-	-	-	-	-	-	6
BF I	-	-	-	-	-	-	6
BF II	-	-	-	**1**	-	-	6
CK	-	-	-	-	-	-	6
CF I	-	-	-	-	**1**	-	6
CF II	-	-	-	-	-	-	6
DK	-	-	-	-	-	-	6
DF I	-	-	-	-	-	-	6
DF II	-	-	-	**1**	-	-	6

Notes: Incubation time glucose dishes: 76.5 h Incubation time rifampicin dishes: 96 h
Four mutant colonies from four different dishes samples were cultivated in LB media containing 30 mg/ml rifampicin media for 7 days.

Assay number: 2019/20

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0,616	3,08	d	0,555	2,78
b	0,566	2,83	e	0,610	3,05
c	0,552	2,76	f	0,547	2,74

Cultures chosen for further analysis: a=A, b=B, d=C, e=D

Mutagenesis: 45.5 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	20	4	3	-	1
AK II	44	15	6	-	1
AK III	23	3	-	1	-
BK I	36	6	34	-	3
BK II	62	3	10	-	1
BK III	19	3	6	-	2
CK I	28	2	9	-	-
CK II	50	7	4	-	-
CK III	14	2	2	-	1
DK I	43	5	15	-	1
DK II	53	12	1	5	-
DK III	45	4	2	1	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	**1**	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	**2**	-	**1**	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Notes: Incubation time glucose dishes: 70.5 h Incubation time rifampicin dishes: 96 h
Four mutant colonies from four different dishes samples were cultivated in LBmedia containing 30 mg/ml rifampicin media for 6 days.

Assay number: 2019/21

Bacterial strain: AB 1884 (*uvrC*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.644	3.22	d	0.653	3.27
b	0.655	3.27	e	0.656	3.28
c	0.665	3.33	f	0.657	3.28

Cultures chosen for further analysis: a=A, b=B, d=C, e=D

Mutagenesis: 45.5 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	178	29	15	-	-
AK II	221	31	31	-	1
AF	166	5	6	1	-
BK I	356	37	13	1	2
BK II	216	31	23	1	-
BF	78	12	6	-	-
CK I	183	24	7	1	1
CK II	198	17	5	-	-
CF	100	5	19	-	-
DK I	301	24	21	3	5
DK II	83	9	7	-	-
DF	223	20	17	2	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	**1**	-	**2**	-	-	6
AF	-	**2**	-	-	-	**1**	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BF	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CF	-	**1**	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	**1**	-	-	-	-	-	6
DF	-	**1**	-	-	-	-	6

Notes: Heating cabinet stopped functioning sometime between 26/4-19 14:00 and 27/4-19 11:45. Plates were moved when the malfunction was discovered
Incubation time glucose dishes: 68.5 h Incubation time rifampicin dishes: 96 h
Four mutant colonies from four different dishes samples were cultivated in LB media containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/22

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.597	2.98	d	0.600	3.00
b	0.597	2.99	e	0.611	3.05
c	0.594	2.97	f	0.591	2.95

Cultures chosen for further analysis: a=A, b=B, c=C, d=D

Mutagenesis: 45.5 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	24	11	17	2	3
AK II	5	1	-	-	-
AK III	21	1	1	-	1
BK I	4	-	4	-	1
BK II	9	1	1	-	-
BK III	5	4	-	-	-
CK I	6	1	2	1	-
CK II	5	-	-	-	1
CK III	1	1	1	-	-
DK I	-	-	-	-	-
DK II	1	-	-	-	-
DK III	-	-	-	-	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	**2**	-	-	-	**1**	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	-	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 93 h Incubation time rifampicin dishes: 96 h
One mutant colonies from one sample was cultivated in LB media containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/23

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.617	3.09	d	0.622	3.11
b	0.635	3.18	e	0.621	3.11
c	0.628	3.14	f	0.625	3.12

Cultures chosen for further analysis: a=A, d=B, e=C, f=D

Mutagenesis: 45.5 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	184	19	19	3	1
AK II	273	31	44	-	3
AK III	80	9		-	1 -
BK I	52	2	9	3	-
BK II	17	4	7	-	1
BK III	51	9	2	-	1
CK I	46	19	12	1	1
CK II	74	13	17	1	-
CK III	57	14	13	2	3
DK I	135	18	16	3	2
DK II	127	22	16	1	2
DK III	142	14	16	1	1

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	**32**	**42**	**41**	**51**	**37**	**48**	6
CK II	-	-	**2**	**1**	**1**	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	-	-	-	6
DK II	-	-	-	-	-	-	6
DK III	-	-	-	-	-	-	6

Notes: Incubation time glucose dishes: 51.5 h Incubation time rifampicin dishes: 96 h
two mutant colonies from two different samples were cultivated in LB media containing 30 mg/ml rifampicin for 7 days.

Assay number: 2019/24

Bacterial strain: AB 1885 (*uvrB*$^-$)

Overnight culture: 21 h

OD Measurements:

Culture nr	OD at 600 nm	Bact/ml (×10^8)	Culture nr	OD at 600 nm	Bact/ml (×10^8)
a	0.606	3.03	d	0.612	3.06
b	0.619	3.10	e	0.618	3.09
c	0.618	3.09	f	0.614	3.07

Cultures chosen for further analysis: a=A, c=B, d=C, f=D

Mutagenesis: 45.5 h

Notes:

Number of colonies on glucose dishes					
Dish	10^{-5}	10^{-6}		10^{-7}	
AK I	4	1	-	-	1
AK II	8	-	3	-	-
AK III	8	-	3	-	-
BK I	13	3	6	-	-
BK II	51	6	16	1	1
BK III	36	2	2	-	-
CK I	3	1	1	-	-
CK II	52	7	7	-	-
CK III	37	4	7	-	-
DK I	16	4	2	1	1
DK II	15	6	4	-	-
DK III	13	9	5	1	-

Number of colonies on rifampicin dishes							**Number of dishes**
AK I	-	-	-	-	-	-	6
AK II	-	-	-	-	-	-	6
AK III	-	-	-	-	-	-	6
BK I	-	-	-	-	-	-	6
BK II	-	-	-	-	-	-	6
BK III	-	-	-	-	-	-	6
CK I	-	**1**	-	-	-	-	6
CK II	-	-	-	-	-	-	6
CK III	-	-	-	-	-	-	6
DK I	-	-	-	**2**	**2**	-	6
DK II	-	-	-	-	-	-	6
DK III	**1**	-	**1**	-	-	-	6

Notes: Incubation time glucose dishes: 92 h Incubation time rifampicin dishes: 94 h
Three mutant colonies from three different samples were cultivated in LB media containing 30 mg/ml rifampicin for 7 days.

A.1.4 Results, mutation rate calculations and BLAST diagrams and chromatograms of mutant colonies

Table A.1.17: Complete list of results

Assay	Sample	Mutation	
uvrB⁻			
2019/2	AK	C→T	GC→AT
	BF II	A→T	TA→AT
	CF I	Mutation occurred outside of sequenced area	
2019/3	BF I	C→A	GC→TA
2019/4	AK	G→T	CG→AT
	CK	C→A	GC→TA
	DF II	C→G	GC→CG
2019/6	BF I	C→T	GC→AT
	BF II	G→T	CG→AT
	CK	G→T	CG→AT
	CF I	C→T	GC→AT
	CF II	C→T	GC→AT
	DK	G→A	CG→TA
	DF I	C→T	GC→AT
	DF III	C→T	GC→AT
2019/8	AF I	C→T	GC→AT
	BF I	Mutation occurred outside of sequenced area	
	CK	C→A	GC→TA
	CF I	Mutation occurred outside of sequenced area	
	CF II	Mutation occurred outside of sequenced area	
	DK	Mutation occurred outside of sequenced area	
2019/10	AK	G→T	CG→AT
	AF I	C→T	GC→AT
	AF II	A→G	TA→CG
	BK	G→A	CG→TA
	BF I	C→T	GC→AT
	BF II	G→A	CG→TA
	CK	C→A	GC→TA
	CF I	No match when run through BLAST program	
	CF II	Too poor quality to determine mutation	
	DF I	C→T	GC→AT
	DF II	G→A	CG→TA
2019/11	AF I	Too poor quality to determine mutation	
	AF II	C→T	GC→AT
	BK	C→A	GC→TA
	BF I	C→T	GC→AT
	DF II	C→T	GC→AT
2019/15	AK I	Too poor quality to determine mutation	
	AK I*	C→A	GC→TA
	AK III*	G→C	CG→GC
	BK I*	C→A	GC→TA
	BK II*	C→G	GC→CG
2019/18	AK I	G→T	CG→AT
	CK I	C→T	GC→AT
2019/20	AK II	No match when run through BLAST program	
	BK I	C→T	GC→AT

2019/23	CK I	G→T	CG→AT
	CK III	C→T	GC→AT
2019/24	CK I	C→T	GC→AT
	DK I	C→T	GC→AT
	DK II	No match when run through BLAST program	
2016/1	AK I	C→T	GC→AT
	AK II	C→A	GC→TA
2016/2	AK I	No match when run through the BLAST program	
	AK III	C→T	GC→AT
	AK IV	C→A	GC→TA
	AK V	G→T	CG→AT
2016/3	AK III	C→T	GC→AT
	AK V	C→T	GC→AT
	AF I	C→T	GC→AT
	AF II	G→A	CG→TA
	AF III	C→T	GC→AT
	AF IV	C→T	GC→AT
	AF V	C→T	GC→AT
	BK II	Too poor quality to determine mutation	
	BK III	C→T	GC→AT
	BK IV	T→C	AT→GC
	BK V	C→T	GC→AT
	BF I	C→T	GC→AT
	BF II	Too poor quality to determine mutation	
	BF III	C→T	GC→AT
	BF IV	Too poor quality to determine mutation	
	BF V	Too poor quality to determine mutation	
2016/4	AK III	Too poor quality to determine mutation	
	AK IV	C→T	GC→AT
	AF I	G→A	CG→TA
	AF II	Too poor quality to determine mutation	
	AF III	Too poor quality to determine mutation	
	AF IV	C→T	GC→AT
	AF V	No match when run through BLAST program	
	BK I	Too poor quality to determine mutation	
	BK II	C→T	GC→AT
	BK III	A→T	TA→AT
	BK IV	C→T	GC→AT
	BK V	C→T	GC→AT
	BF II	C→T	GC→AT
	BF III	A→T	TA→AT
	BF V	C→A	GC→TA
uvrC^{-}			
2019/19	AF II	G→A	CG→TA
	BF II	A→G	TA→CG
	CF I	No match when run through BLAST program	
	DF II	C→A	GC→TA
2019/21	AK II	C→T	GC→AT
	AF	G→T	CG→AT
	CF	G→T	CG→AT
	DK II	C→T	GC→AT
	DF	C→A	GC→TA

uvrB⁻, spontaneous mutations

33 experiments selected from totally 194 experiments, N_t= 0,5 – 1,5 × 10^9

Culture	N_0	Time (h)	N_t	Mutants/plate						z	r	N_0/N_t
				1	2	3	4	5	6			
2015/72, AK	9000	48	1328000000	0	0	0	0	0	0	0,30	0	0,00000678
2015/72, BK	9000	48	1420000000	0	1	0	0	1	0	0,30	7	0,00000634
2015/72, DK	9000	48	1139000000	0	0	0	0	0	0	0,30	0	0,00000790
2015/75, AK	9000	45	1560666667	0	0	0	0	0	0	0,30	0	0,00000577
2015/75, BK	9000	45	1899333333	3	0	0	1	2	0	0,30	20	0,00000474
2015/75, CK	9000	45	1334000000	0	0	0	0	0	0	0,30	0	0,00000675
2015/75, DK	9000	45	991333333	0	1	0	0	0	0	0,30	3	0,00000908
2016/3, AKII	9000	45	1104000000	0	0	0	0	0	0	0,30	0	0,00000815
2016/3, AKIII	9000	45	1691000000	0	0	2	2	1	1	0,30	20	0,00000532
2016/3, AKIV	9000	45	1127000000	0	0	0	0	0	0	0,30	0	0,00000799
2016/3, AKV	9000	45	1693500000	0	0	0	0	0	1	0,30	3	0,00000531
2016/3, BKII	9000	45	1025333333	0	4	3	6	4	3	0,30	67	0,00000878
2016/3, BKIII	9000	45	1192666667	0	0	0	0	0	1	0,30	3	0,00000755
2016/3, BKIV	9000	45	940666667	0	0	0	0	0	1	0,30	3	0,00000957
2016/3, BKV	9000	45	1430500000	0	0	0	0	1	1	0,30	7	0,00000629
2016/4, AKIV	9000	45	949333333	1	2	1	5	2	0	0,30	37	0,00000948
2016/B6, BKII	9000	46	1816000000	1	0	0	0	0	0	0,30	3	0,00000496
2016/B6, BKIII	9000	46	922666667	0	0	0	0	0	0	0,30	0	0,00000975
2016/B6, BKV	9000	46	1090000000	0	0	0	0	0	0	0,30	0	0,00000826
2019/6, BK	18000	45	1526000000	0	0	0	0	0	0	0,30	0	0,00001180
2019/6, CK	18000	45	1124000000	0	1	0	0	0	0	0,30	3	0,00001601
2019/6, DK	18000	45	1347333333	3	2	2	3	4	1	0,30	50	0,00001336
2019/8 AK	18000	45,5	1661333333	0	0	0	0	0	0	0,30	0	0,00001083
2019/8, CK	18000	45,5	1864000000	0	0	0	0	0	0	0,30	0	0,00000966
2019/10, AK	18000	47,5	1526000000	0	0	0	1	0	0	0,30	3	0,00001180
2019/10, BK	18000	47,5	1308666667	1	0	1	0	2	0	0,30	13	0,00001375
2019/10, CK	18000	47,5	1050666667	0	0	39	0	2	0	0,30	137	0,00001713
2019/11, AK	18000	45	1066666667	0	0	0	0	0	0	0,30	0	0,00001688
2019/11, BK	18000	45	1006000000	1	0	0	1	1	1	0,30	13	0,00001789
2019/15 AK I	18000	45,5	1037333333	0	0	0	0	0	2	0,30	7	0,00001735
2019/15, AK II	18000	45,5	1384000000	0	0	0	0	0	0	0,30	0	0,00001301
2019/15, BK I	18000	45,5	1061333333	0	0	0	0	7	4	0,30	37	0,00001696
2019/15, BK II	18000	45,5	954666667	0	0	8	0	8	0	0,30	53	0,00001885
2019/15, BK III	18000	45,5	1019333333	0	0	0	0	0	0	0,30	0	0,00001766
Average	12971	46	1282127451									
SD	4536	1	294232292									
Median	9000	45	1165833333									
Average			1,3									
SD			0,3									
Median			1,2									

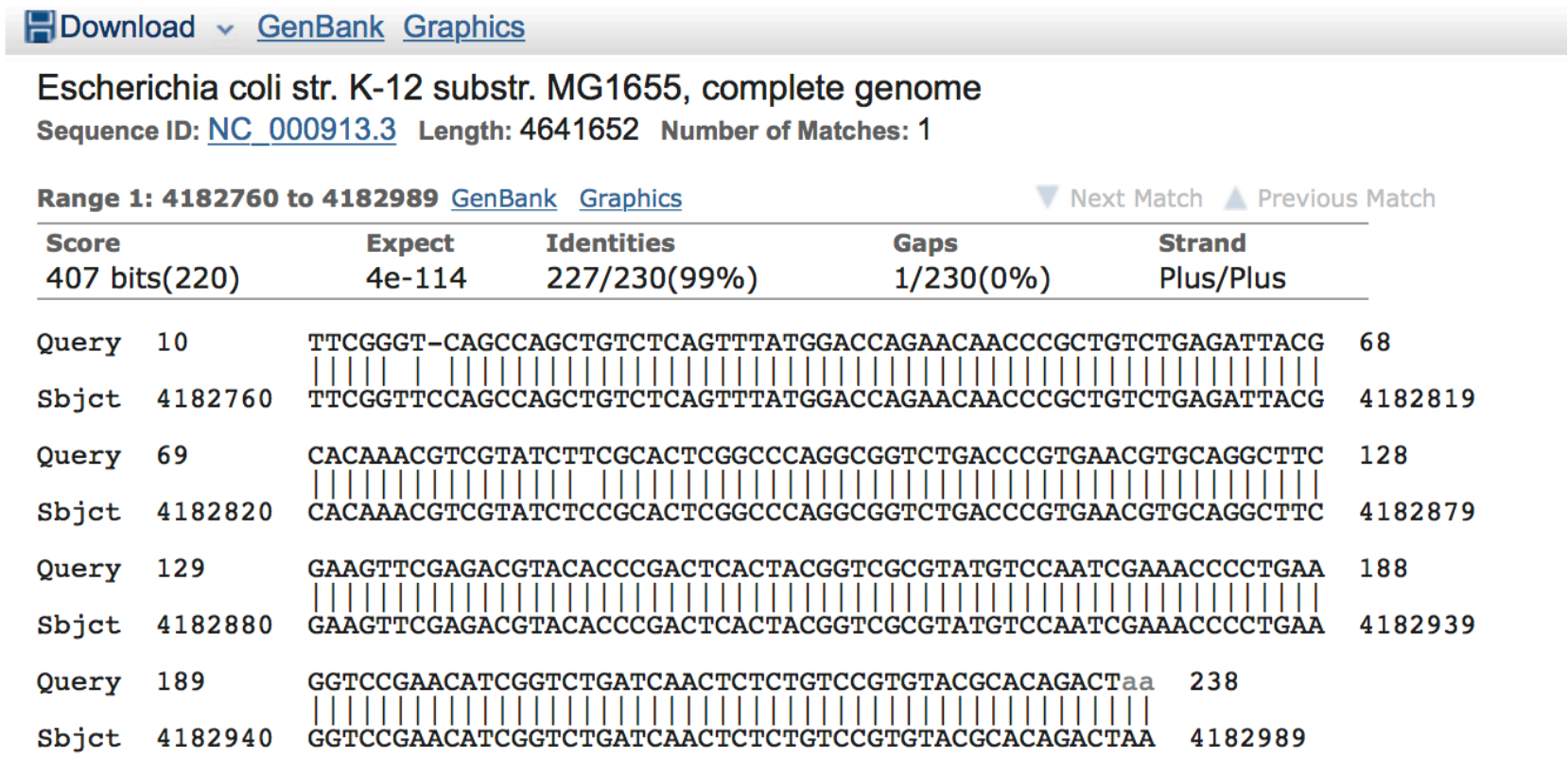
Zeros= 14
C= 33
p_0= 0,424
m_{obs}= 0,86
m_{act}= 1,66
μ=1.42 × 10^{-9}

uvrB⁻, 0.1 mM f⁵dU induced mutations

37 experiments selected from totally 127 experiments, N_t= 0,5 – 1,5 × 10^9

Experiment	N_0	Time (h)	N_t	Mutants/ plate						z	r	N_0/N_t
				1	2	3	4	5	6			
2014/62, AFI	9000		2823200000	10	11	25	18	20	14	0,30	327	0,000003
2015/72, AFI	9000	48	1724500000	0	0	0	0	0	0	0,30	0	0,000005
2015/75, AFI	9000	45	1102500000	3	1	0	1	0	1	0,30	20	0,000008
2015/75, BFI	9000	45	2028500000	0	1	1	1	0	2	0,30	17	0,000004
2015/75, BFII	9000	45	1484000000	5	4	2	3	4	4	0,30	73	0,000006
2015/75, CFI	9000	45	1547000000	0	1	0	0	0	0	0,30	3	0,000006
2015/75, CFII	9000	45	1010500000	0	0	0	0	0	0	0,30	0	0,000009
2015/80, BF	9000	46	1223000000	0	0	0	0	0	0	0,30	0	0,000007
2016/3, AFII	9000	45	991333333	0	0	0	0	0	1	0,30	3	0,000009
2016/3, AFIII	9000	45	1390000000	0	0	2	1	1	1	0,30	17	0,000006
2016/3, AFIV	9000	45	1435000000	0	0	0	1	2	1	0,30	13	0,000006
2016/3, AFV	9000	45	1519500000	0	0	0	0	0	4	0,30	13	0,000006
2016/3, BFI	9000	45	1447333333	0	0	0	0	0	1	0,30	3	0,000006
2016/3, BFII	9000	45	1019333333	0	0	0	0	0	3	0,30	10	0,000009
2016/3, BFIII	9000	45	1085333333	0	0	0	0	2	0	0,30	7	0,000008
2016/3, BFIV	9000	45	1209333333	0	0	0	0	0	1	0,30	3	0,000007
2016/3, BFV	9000	45	1700000000	0	4	3	1	3	4	0,30	50	0,000005
2016/4, AFI	9000	45	1004000000	0	0	1	0	0	1	0,30	7	0,000009
2016/4, AFV	9000	45	1678000000	1	0	0	0	1	0	0,30	7	0,000005
2019/6, BF II	18000	45	1543000000	1	0	2	1	0	1	0,30	17	0,000012
2019/6, CF I	18000	45	1123500000	0	1	1	1	0	0	0,30	10	0,000016
2019/6, CF II	18000	45	1150500000	0	0	0	0	2	2	0,30	13	0,000016
2019/6, DF I	18000	45	1045000000	0	0	2	0	3	2	0,30	23	0,000017
2019/6, DF II	18000	45	932000000	0	1	3	1	1	1	0,30	23	0,000019
2019/8, AF I	18000	45,5	1238500000	2	1	0	1	0	0	0,30	13	0,000015
2019/8, BF I	18000	45,5	1587500000	3	4	3	4	6	5	0,30	83	0,000011
2019/8, CF I	18000	45,5	1118000000	1	0	2	3	0	2	0,30	27	0,000016
2019/8, CF II	18000	45,5	1220000000	0	1	1	0	1	0	0,30	10	0,000015
2019/10, AF I	18000	47,5	1538500000	0	1	0	1	1	2	0,30	17	0,000012
2019/10, AF II	18000	47,5	1173000000	1	2	3	12	0	1	0,30	63	0,000015
2019/10, BF II	18000	47,5	1425500000	0	3	0	2	1	1	0,30	23	0,000013
2019/10, CF I	18000	47,5	1408000000	0	0	0	0	0	1	0,30	3	0,000013
2019/10, CF II	18000	47,5	1252000000	1	0	0	0	0	0	0,30	3	0,000014
2019/10, DF II	18000	47,5	945500000	1	0	0	0	0	0	0,30	3	0,000019
2019/11, AF I	18000	45	1161000000	0	0	0	0	2	0	0,30	7	0,000016
2019/11, AF II	18000	45	1217500000	0	0	0	0	3	4	0,30	23	0,000015
2019/11, BF I	18000	45	1165000000	1	1	0	0	0	0	0,30	7	0,000015
Avrerage	13378	46	1342334234									
SD	4560	1	358047513									
Median	9000	45	1223000000									
Average			1,3 × 10^9									

		Zeros= 3 C= 37
SD	$0{,}4 \times 10^9$	p_0= 0,081 m_{obs}= 2,51
Median	$1{,}2 \times 10^9$	m_{act}= 4,87 $\mu = 3{,}98 \times 10^{-9}$



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182760 to 4182989 GenBank Graphics Next Match Previous Match

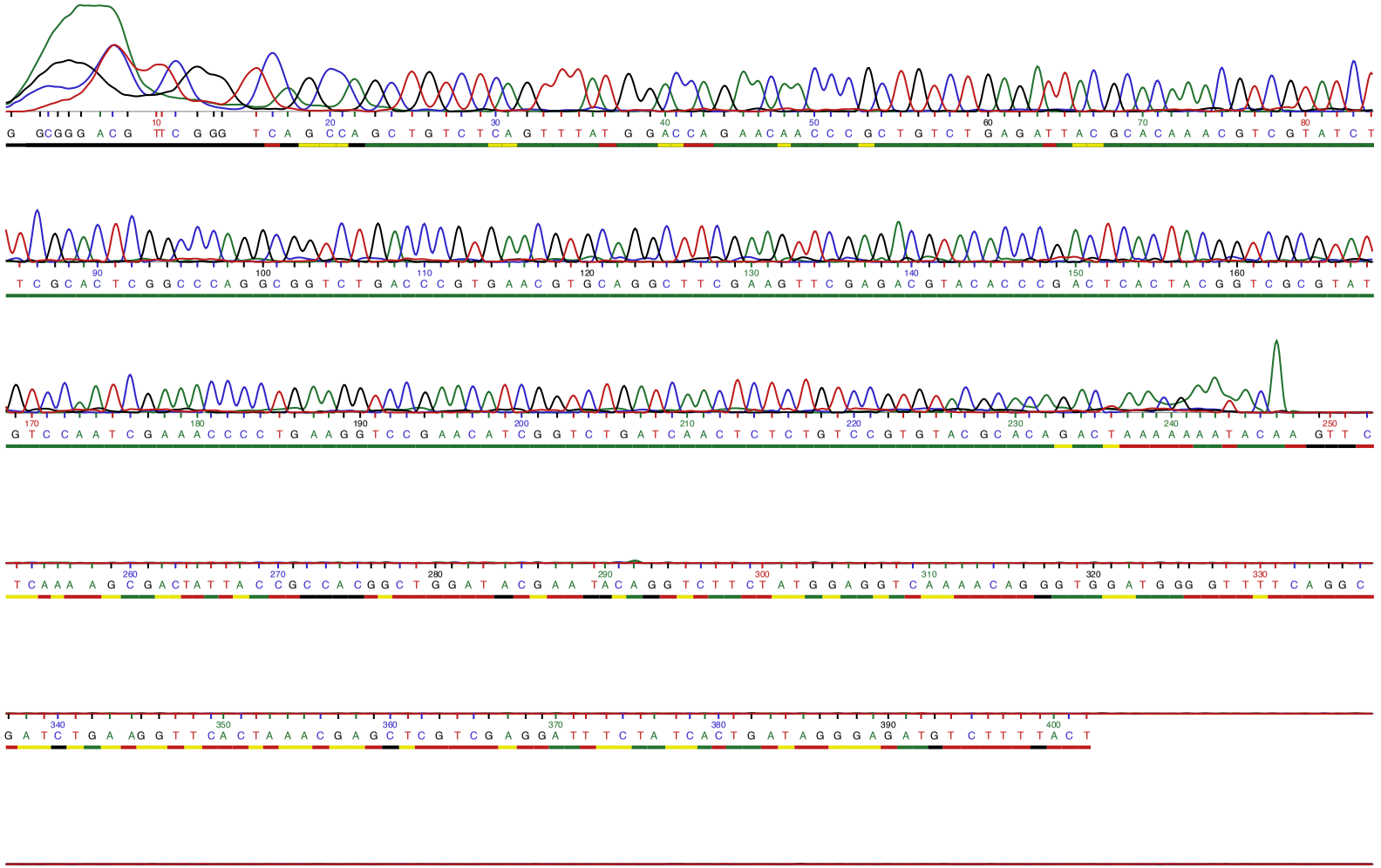
Score	Expect	Identities	Gaps	Strand
407 bits(220)	4e-114	227/230(99%)	1/230(0%)	Plus/Plus

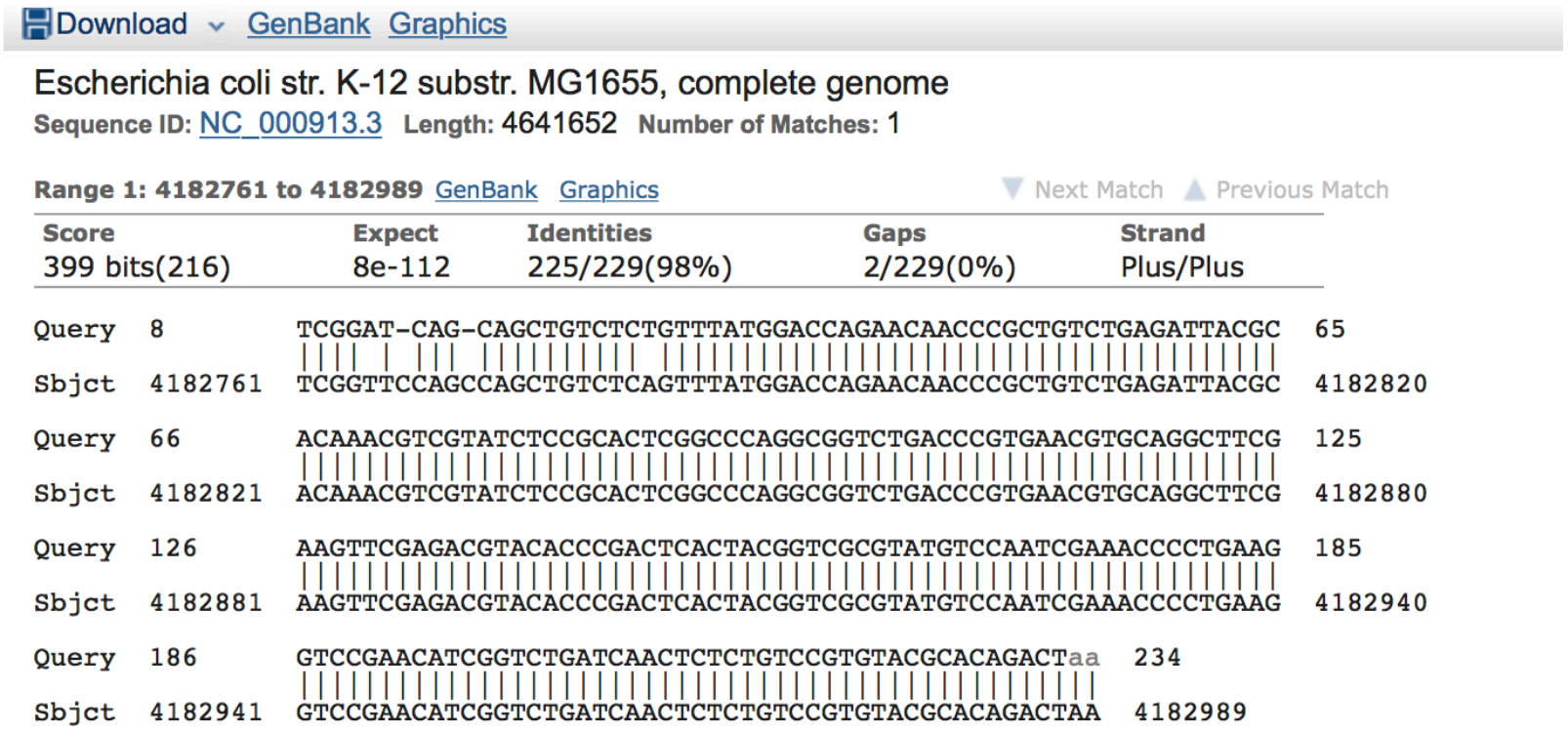
```
Query  10       TTCGGGT-CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  68
                ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182760  TTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  4182819

Query  69       CACAAACGTCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  128
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182820  CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  4182879

Query  129      GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182880  GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  4182939

Query  189      GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  238
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182940  GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/2 AK

Sequence: 50401352

Samples:	16303	Quality:	0 - 9
Bases:	403		10 - 19
Average spacing:	41.0		20 - 29
Average quality >=	10: 81, 20: 55, 30: 232		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182989 GenBank Graphics Next Match Previous Match

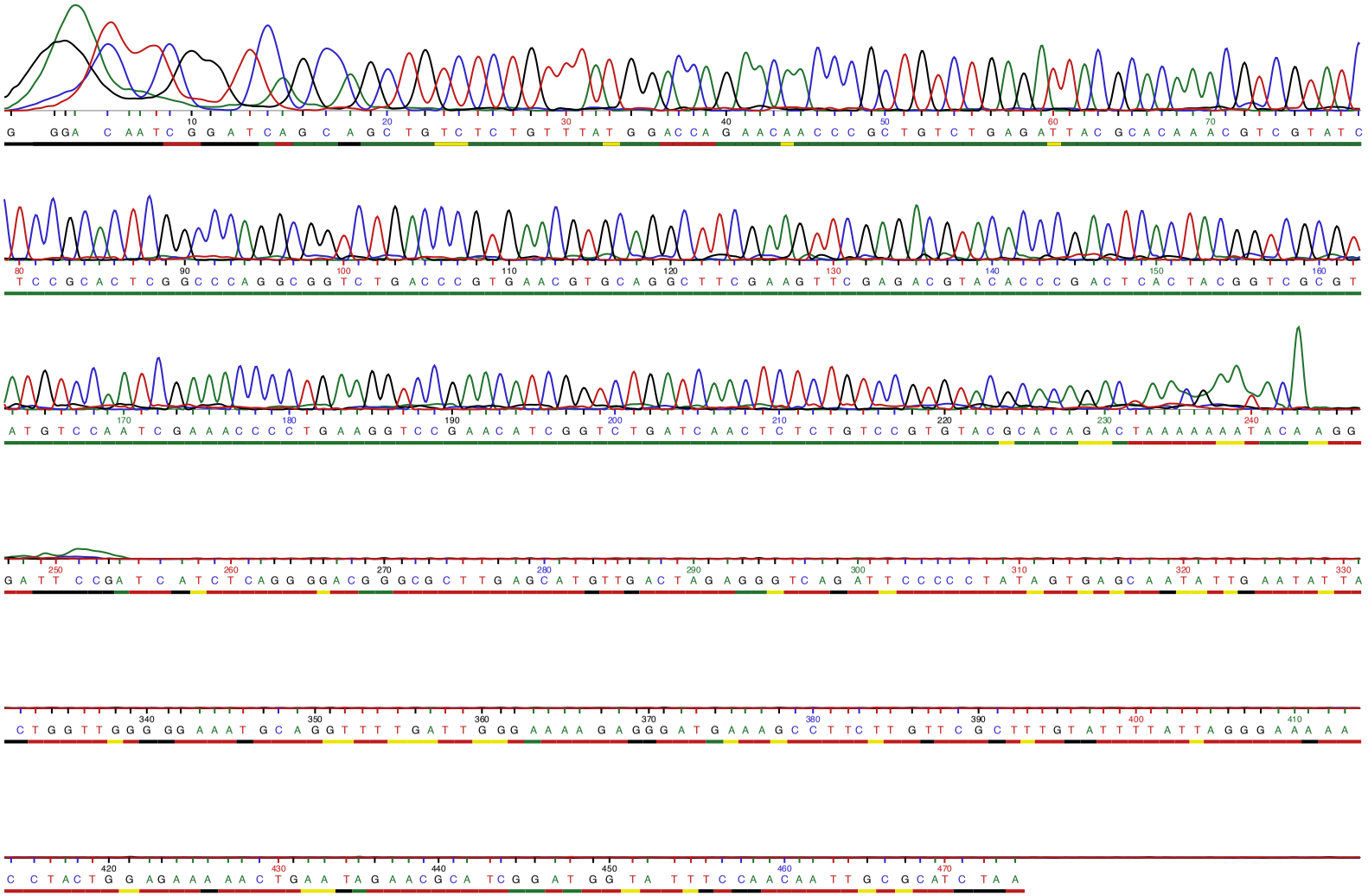
Score	Expect	Identities	Gaps	Strand
399 bits(216)	8e-112	225/229(98%)	2/229(0%)	Plus/Plus

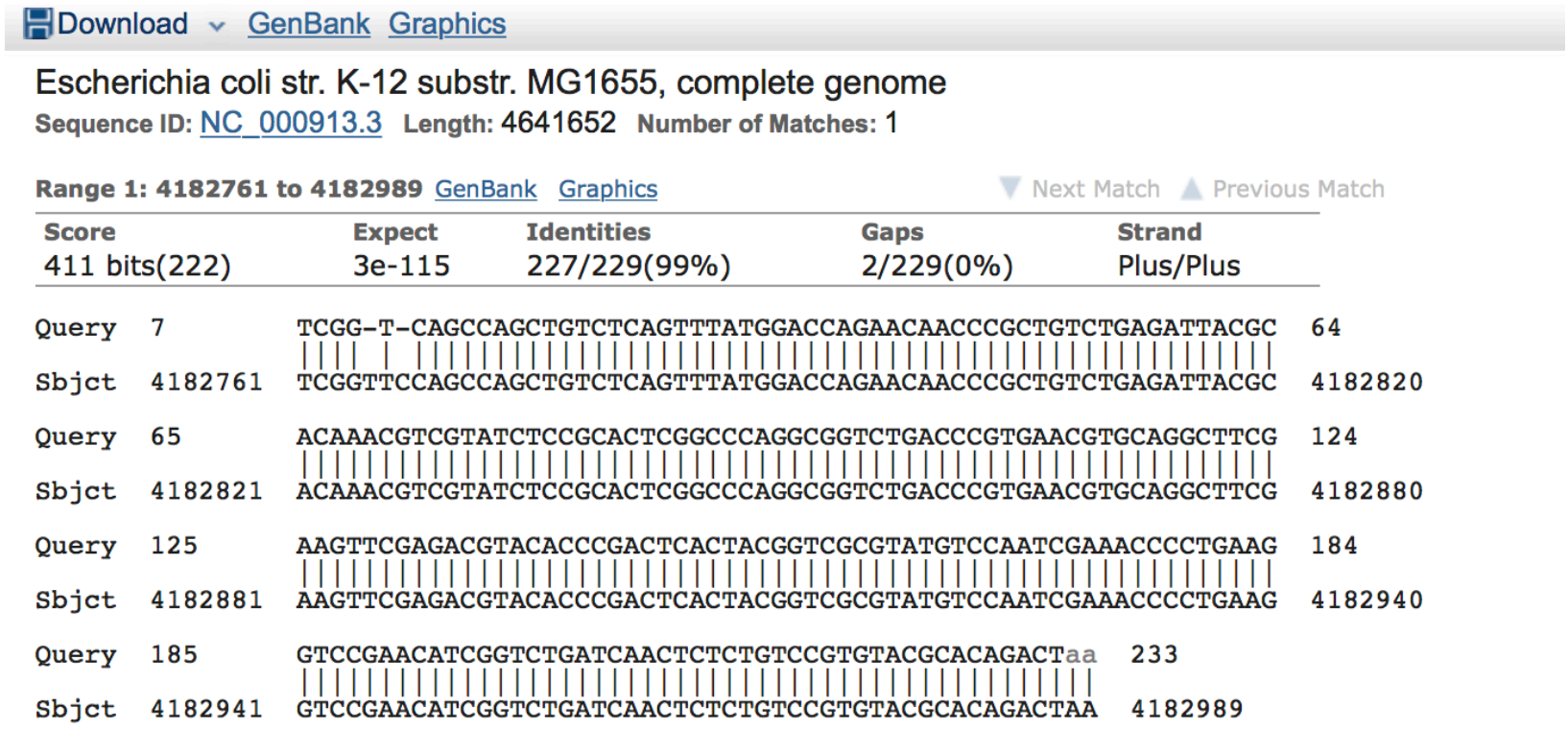
```
Query  8        TCGGAT-CAG-CAGCTGTCTCTGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  65
               ||||  |  |||  |||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  66       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  125
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  126      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  185
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  186      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  234
                ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/2 BF II

Sequence: 50401369

Samples:	16302	Quality:	0 - 9
Bases:	475		10 - 19
Average spacing:	35.0		20 - 29
Average quality >=	10: 173, 20: 42, 30: 218		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182989 GenBank Graphics Next Match Previous Match

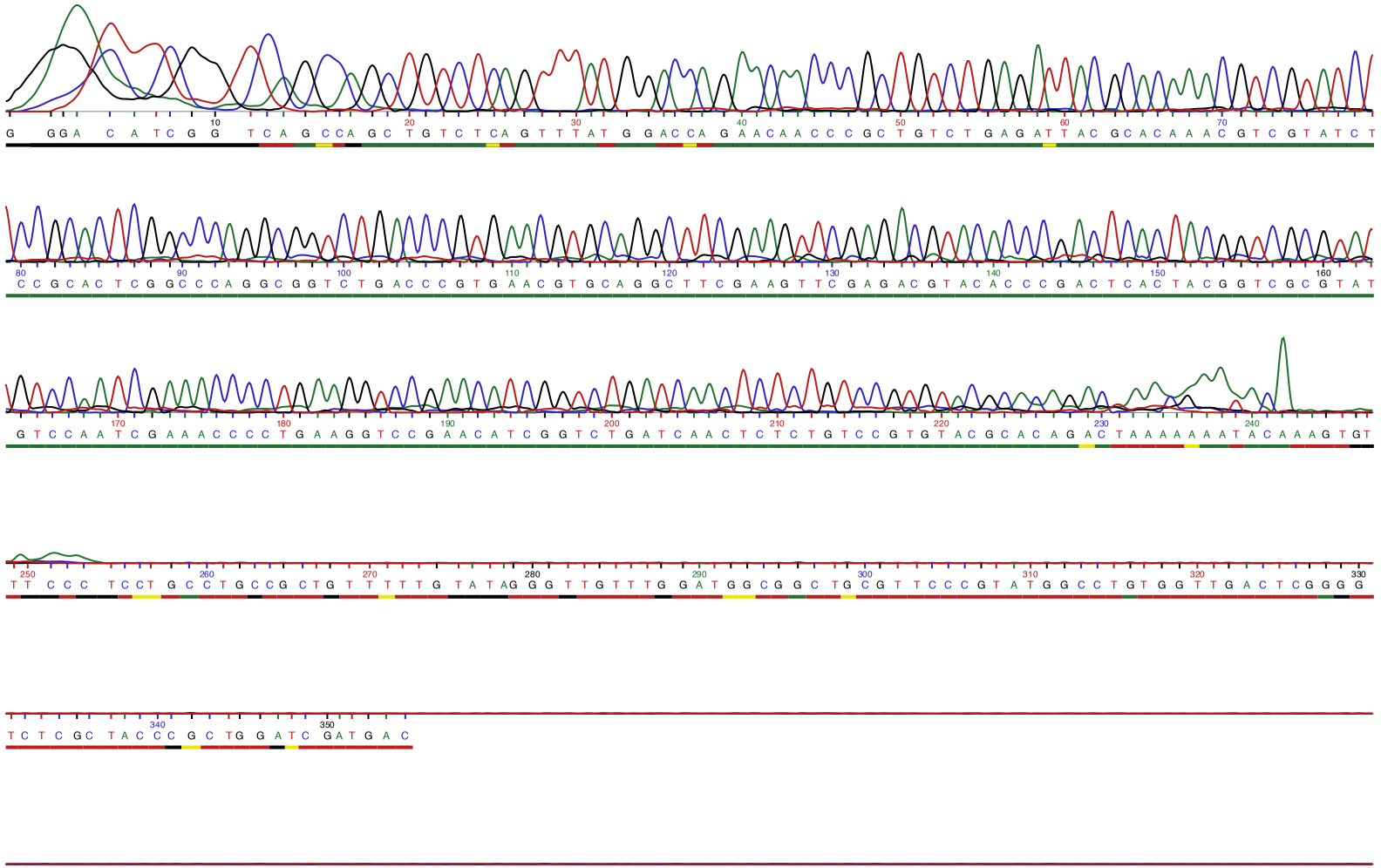
Score	Expect	Identities	Gaps	Strand
411 bits(222)	3e-115	227/229(99%)	2/229(0%)	Plus/Plus

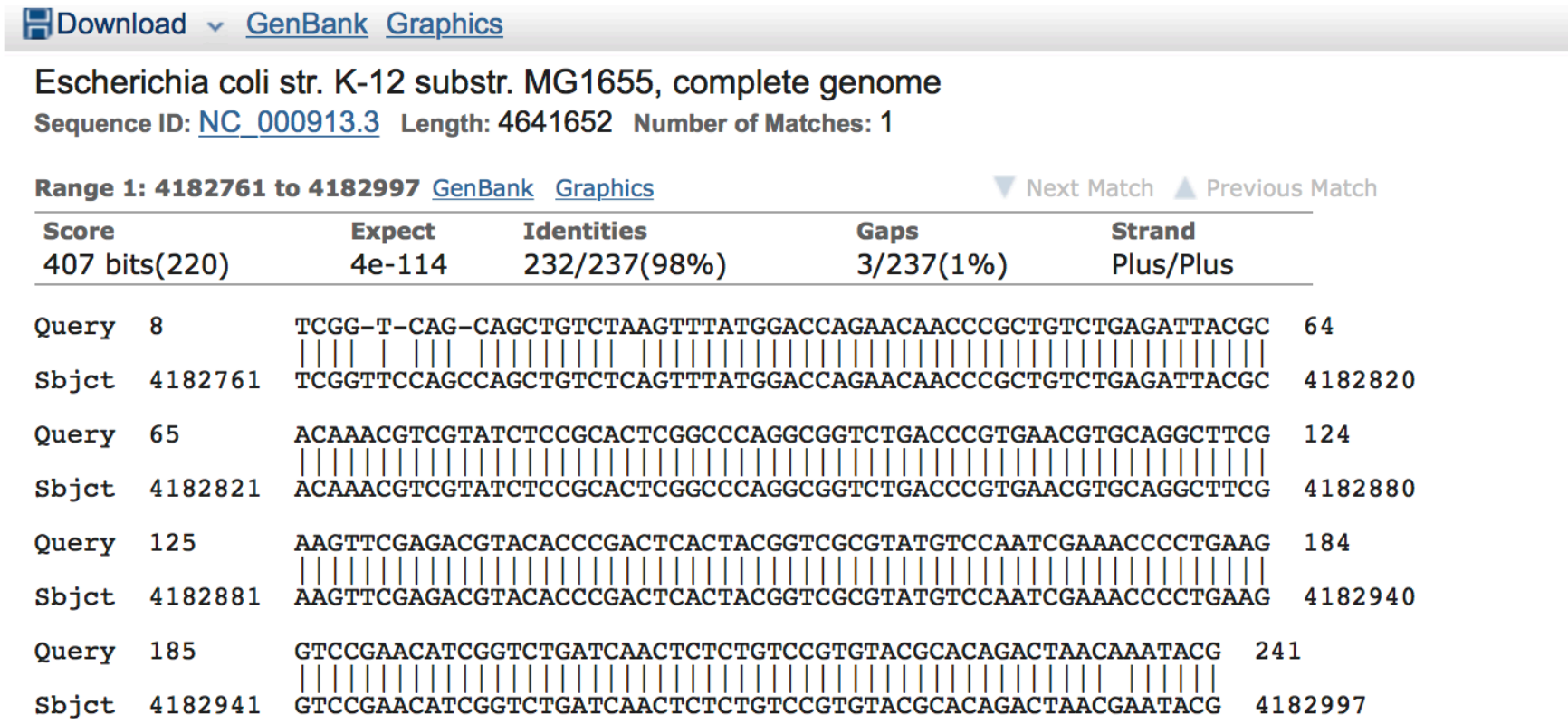
```
Query  7        TCGG-T-CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  64
               |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  65       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  125      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  185      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  233
               |||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/2 CF I

Sequence: 50401376

Samples:	16304	Quality:	0 - 9
Bases:	356		10 - 19
Average spacing:	46.0		20 - 29
Average quality >=	10: 99, 20: 14, 30: 214		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
407 bits(220)	4e-114	232/237(98%)	3/237(1%)	Plus/Plus

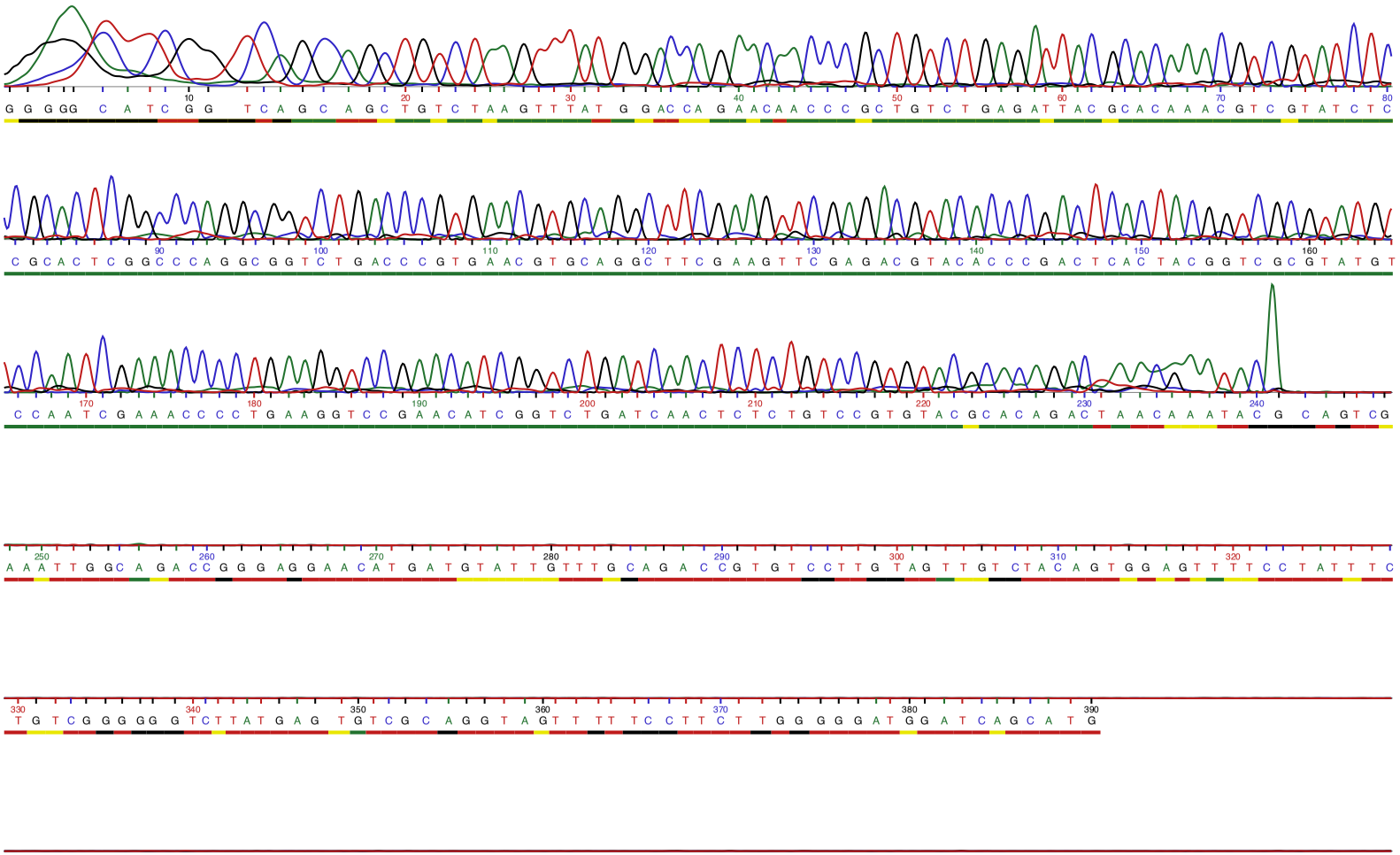
```
Query  8        TCGG-T-CAG-CAGCTGTCTAAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  64
               ||||| | ||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  65       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  125      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  185      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATACG  241
               ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

2019/3 BF I

Sequence: 50401383

Samples:	16208	Quality:	0 - 9
Bases:	391		10 - 19
Average spacing:	42.0		20 - 29
Average quality >=	10: 112, 20: 41, 30: 203		>= 30

12.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182996 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
411 bits(222)	2e-115	227/229(99%)	1/229(0%)	Plus/Plus

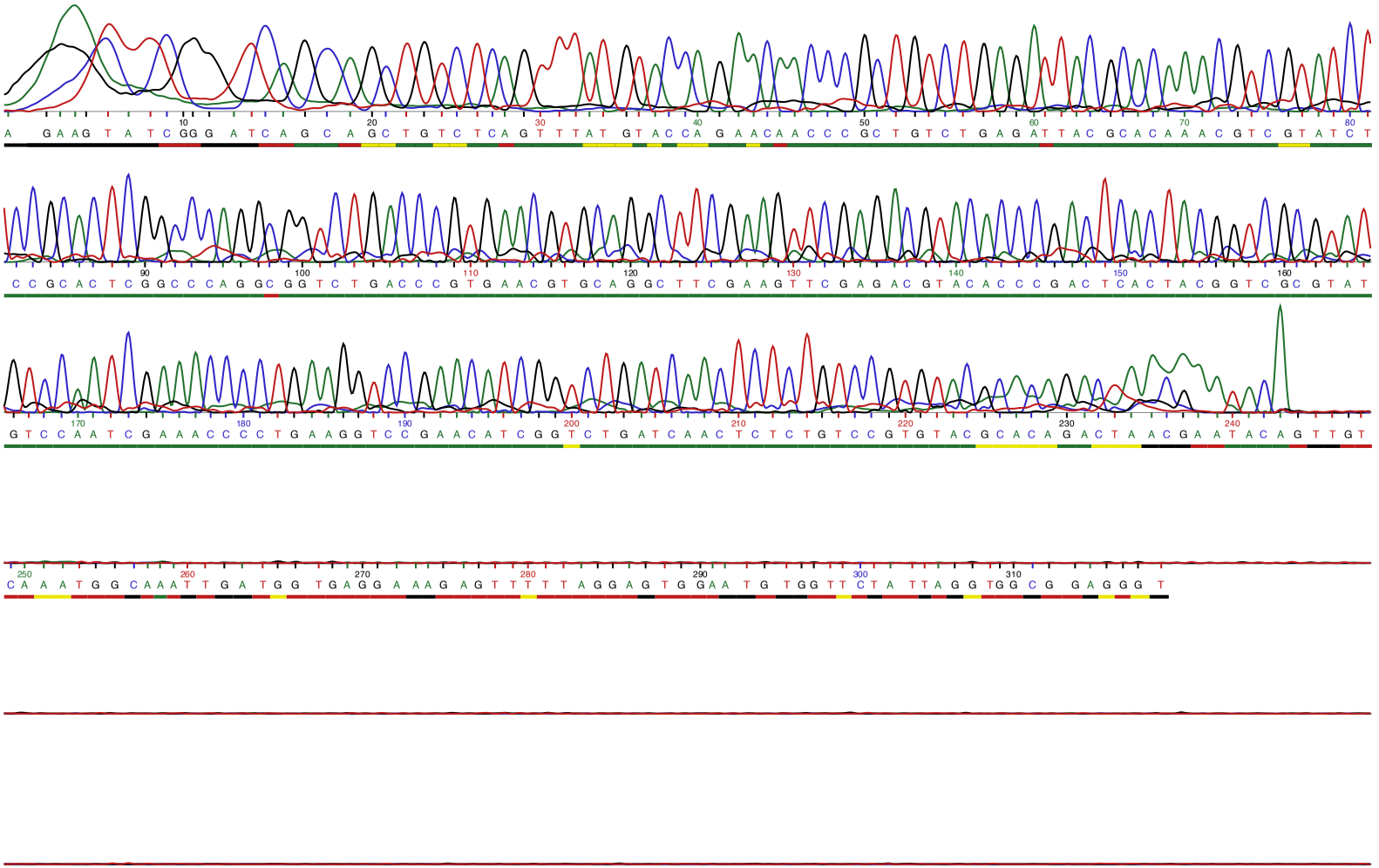
```
Query  15       CAG-CAGCTGTCTCAGTTTATGTACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  73
               ||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  74       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  193
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  242
               ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/4 AK

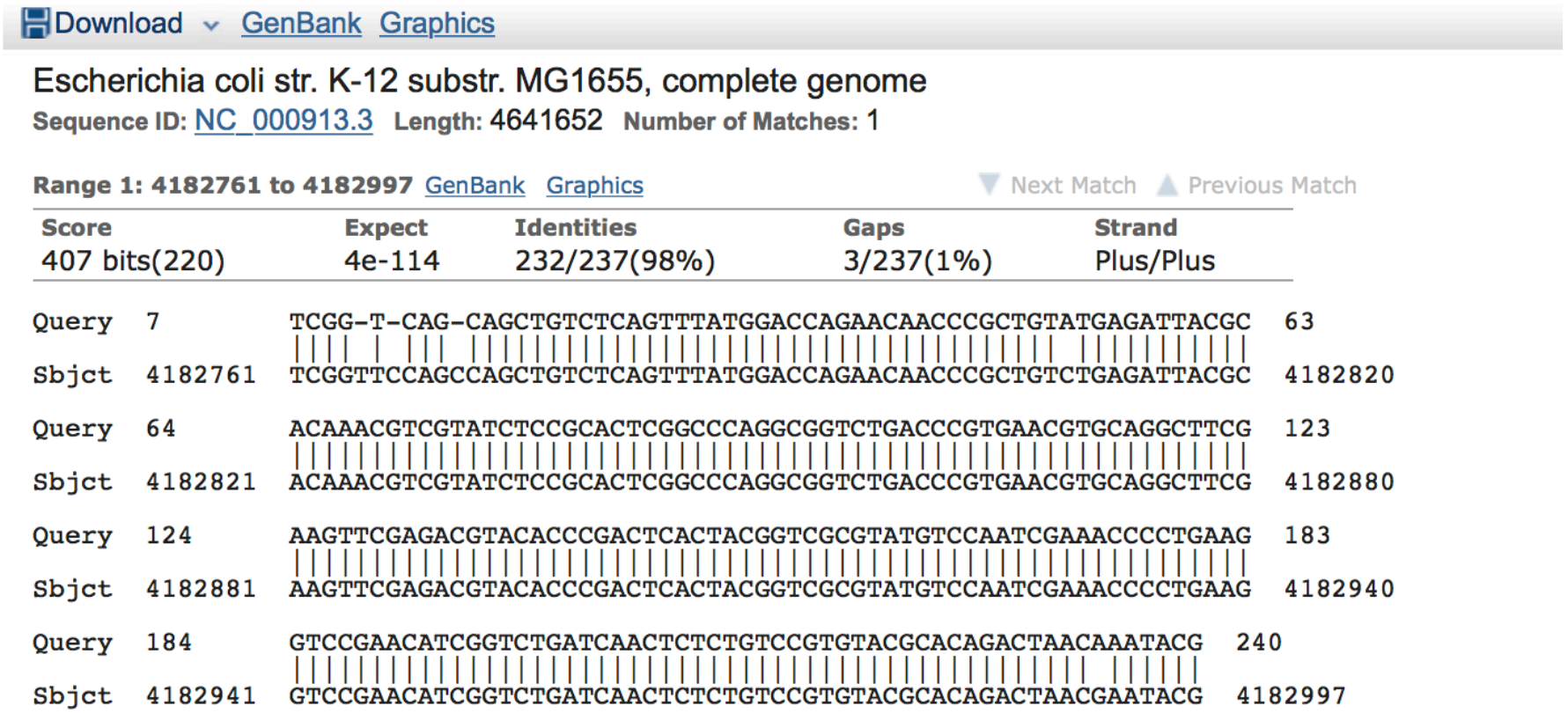
Sequence: 50401390

Samples:	16301	Quality:	0 - 9
Bases:	319		10 - 19
Average spacing:	52.0		20 - 29
Average quality >=	10: 59, 20: 30, 30: 196		>= 30

12.03.2019



2019/4 CK

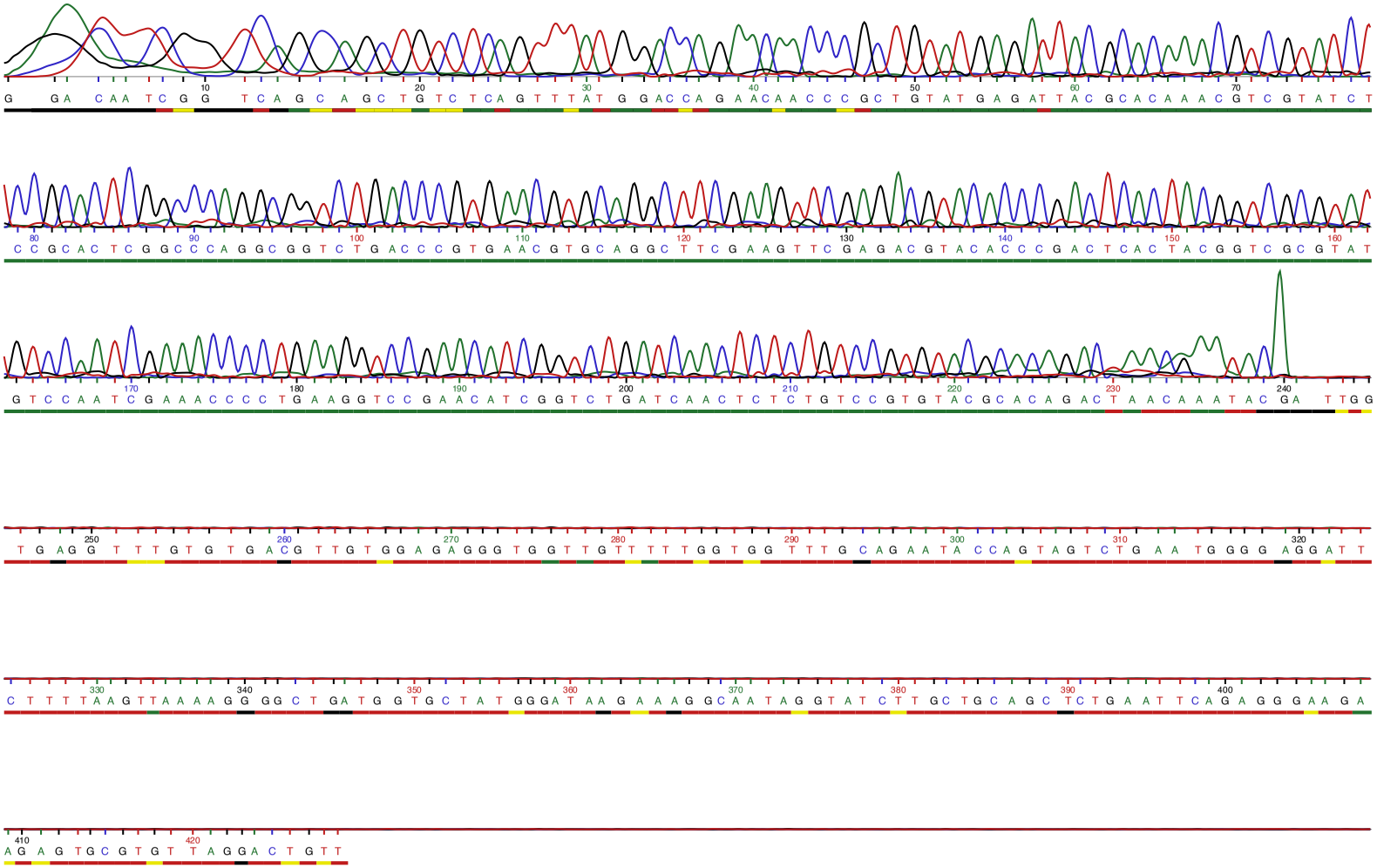

Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182761 to 4182997 GenBank Graphics
Next Match Previous Match
Score 407 bits(220)
Expect 4e-114
Identities 232/237(98%)
Gaps 3/237(1%)
Strand Plus/Plus
Query 7 TCGG-T-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTATGAGATTACGC 63
Sbjct 4182761 TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC 4182820
Query 64 ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG 123
Sbjct 4182821 ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG 4182880
Query 124 AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG 183
Sbjct 4182881 AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG 4182940
Query 184 GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATACG 240
Sbjct 4182941 GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG 4182997

2019/4 CK

Sequence: 50401406

Samples:	16301	Quality:	0 - 9
Bases:	430		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 168, 20: 31, 30: 206		>= 30

12.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
414 bits(224)	3e-116	233/237(98%)	2/237(0%)	Plus/Plus

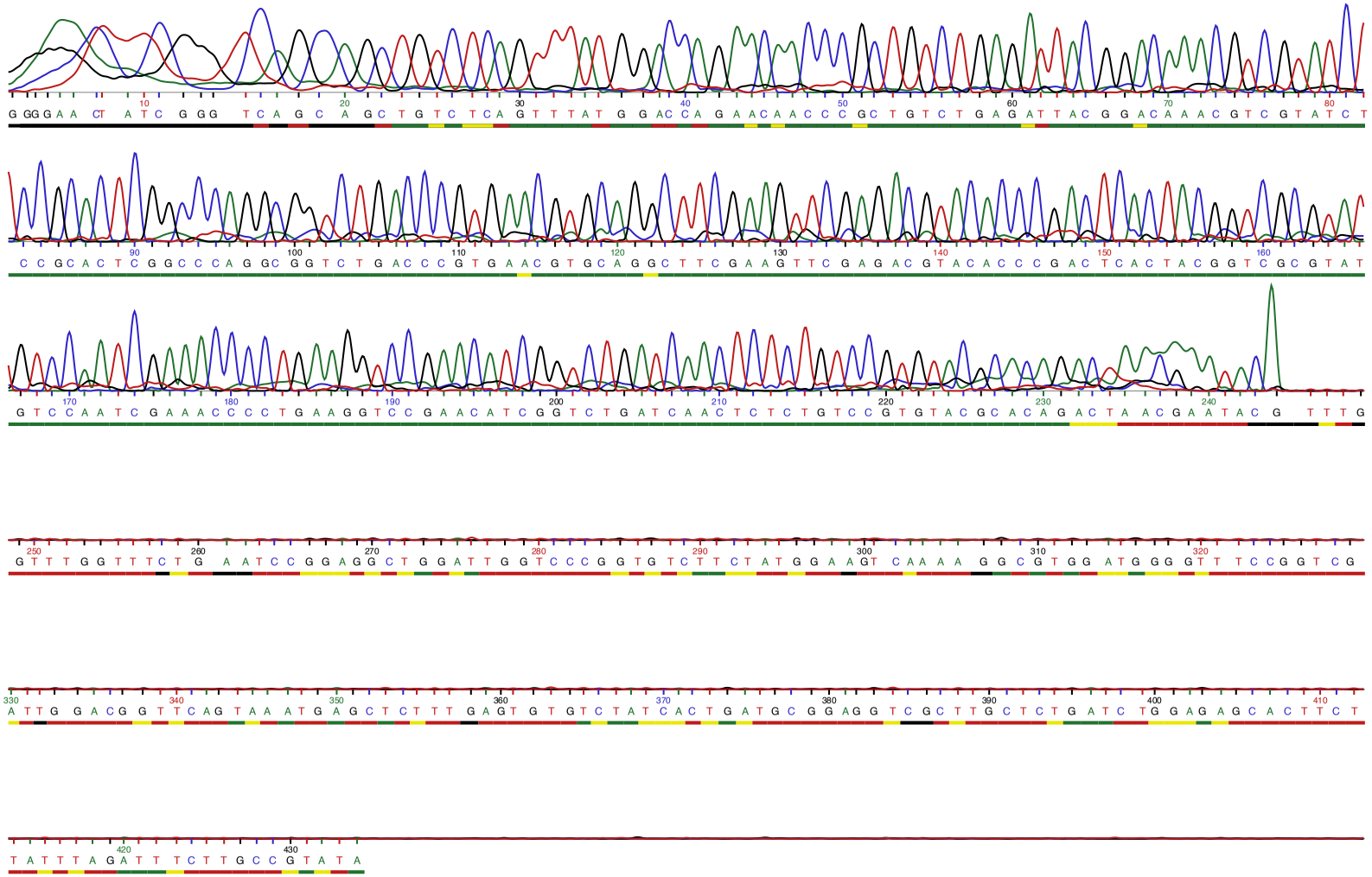
```
Query  10       TCGGGT-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGG  67
                ||||| | ||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  68       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  127
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  128      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  187
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  188      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  244
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

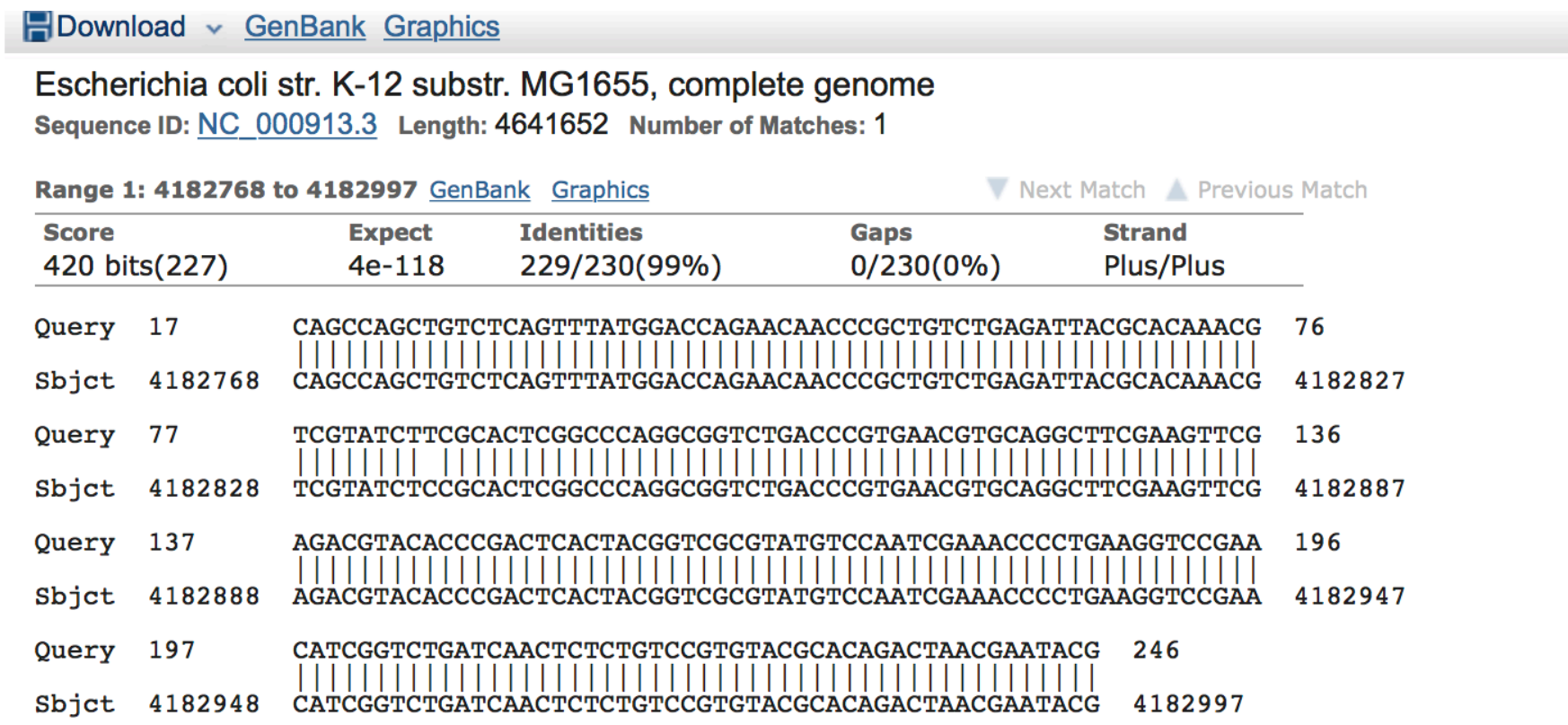
2019/4 DF II

Sequence: 50401413

Samples:	16302	Quality:	0 - 9
Bases:	435		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 129, 20: 56, 30: 218		>= 30

12.03.2019

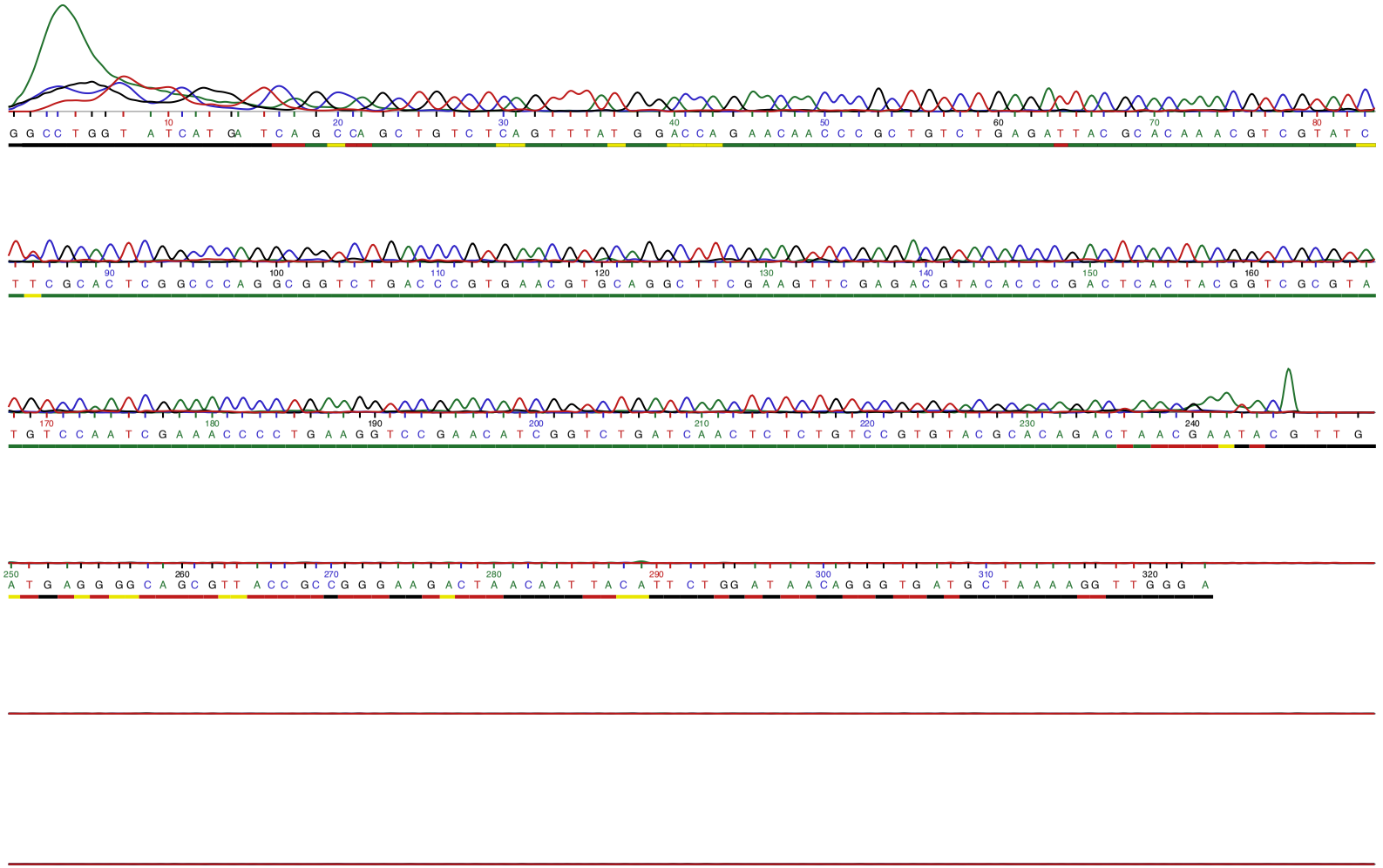


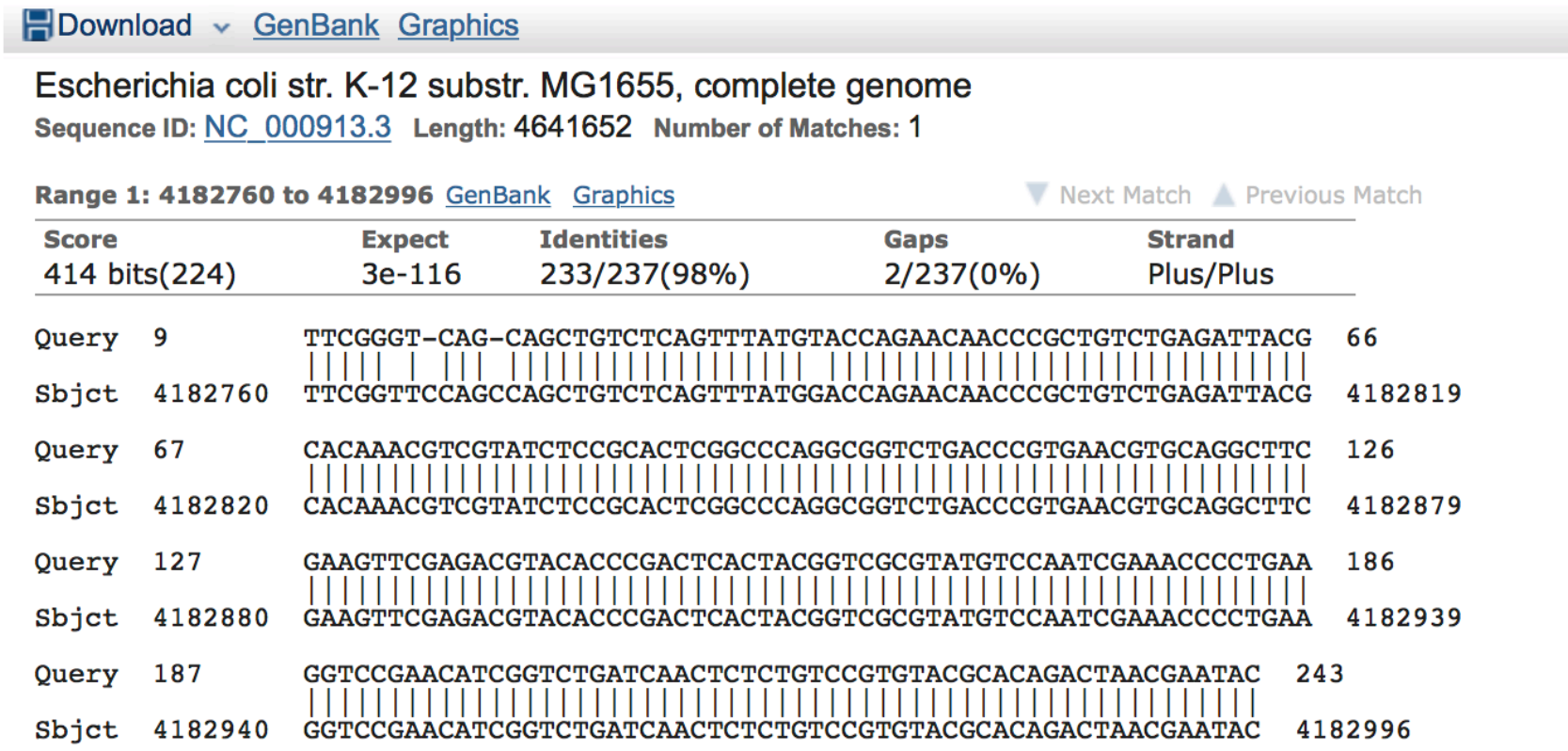

Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182768 to 4182997 GenBank Graphics
Next Match Previous Match
Score
420 bits(227)
Expect
4e-118
Identities
229/230(99%)
Gaps
0/230(0%)
Strand
Plus/Plus
Query 17 CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 76
Sbjct 4182768 CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 4182827
Query 77 TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 136
Sbjct 4182828 TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 4182887
Query 137 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 196
Sbjct 4182888 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 4182947
Query 197 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG 246
Sbjct 4182948 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG 4182997

2019/6 BF I

Sequence: 50401420

Samples:	16301	Quality:	0 - 9
Bases:	324		10 - 19
Average spacing:	51.0		20 - 29
Average quality >=	10: 44, 20: 20, 30: 205		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182760 to 4182996 GenBank Graphics Next Match Previous Match

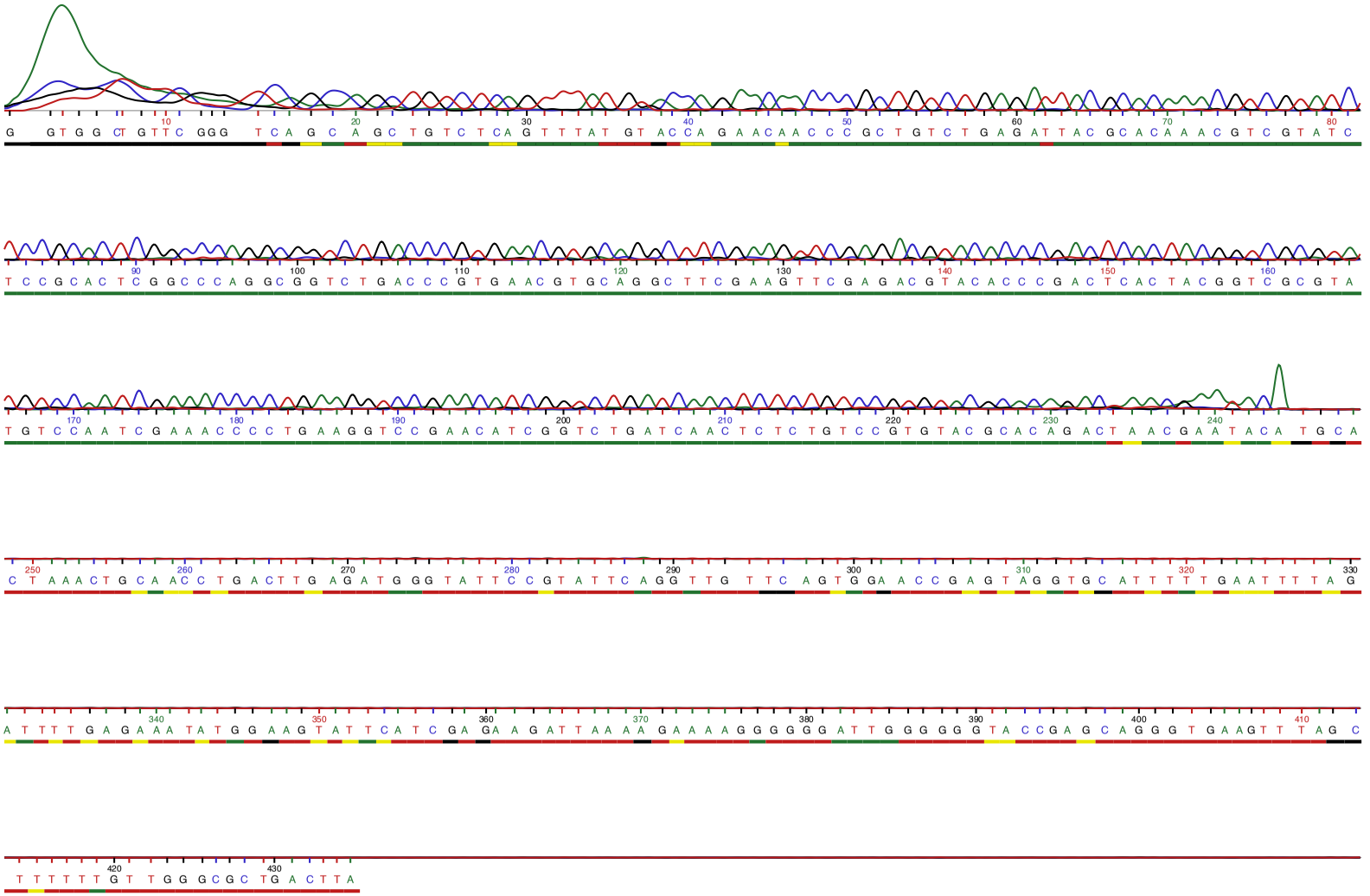
Score	Expect	Identities	Gaps	Strand
414 bits(224)	3e-116	233/237(98%)	2/237(0%)	Plus/Plus

```
Query  9        TTCGGGT-CAG-CAGCTGTCTCAGTTTATGTACCAGAACAACCCGCTGTCTGAGATTACG  66
               ||||| | ||| ||||||||||||||||||| |||||||||||||||||||||||||||||
Sbjct  4182760  TTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  4182819

Query  67       CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182820  CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  4182879

Query  127      GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182880  GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  4182939

Query  187      GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  243
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182940  GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

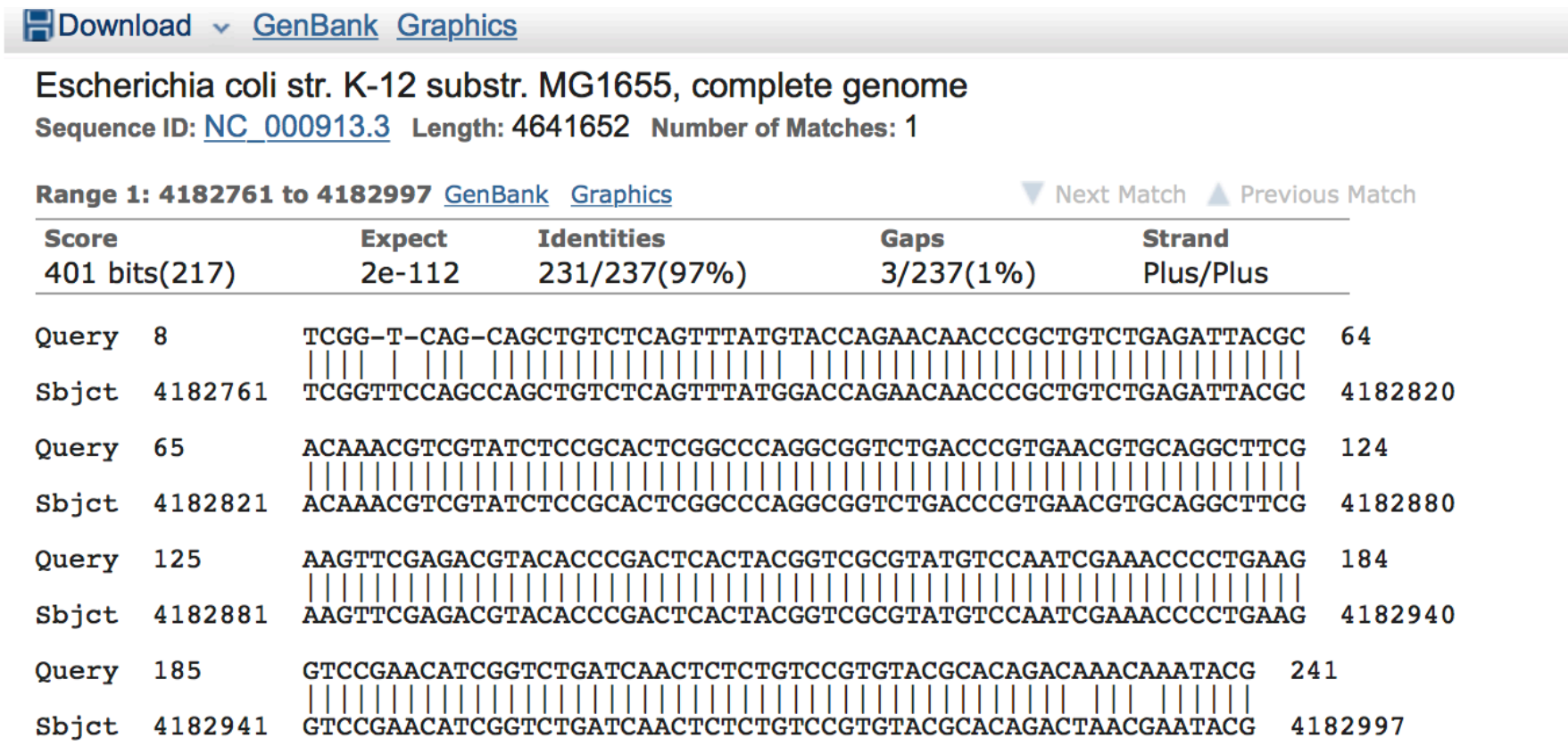
2019/6 BF II

Sequence: 50401437

Samples:	16303	Quality:	0 - 9
Bases:	436		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 141, 20: 41, 30: 224		>= 30



2019/6 CK

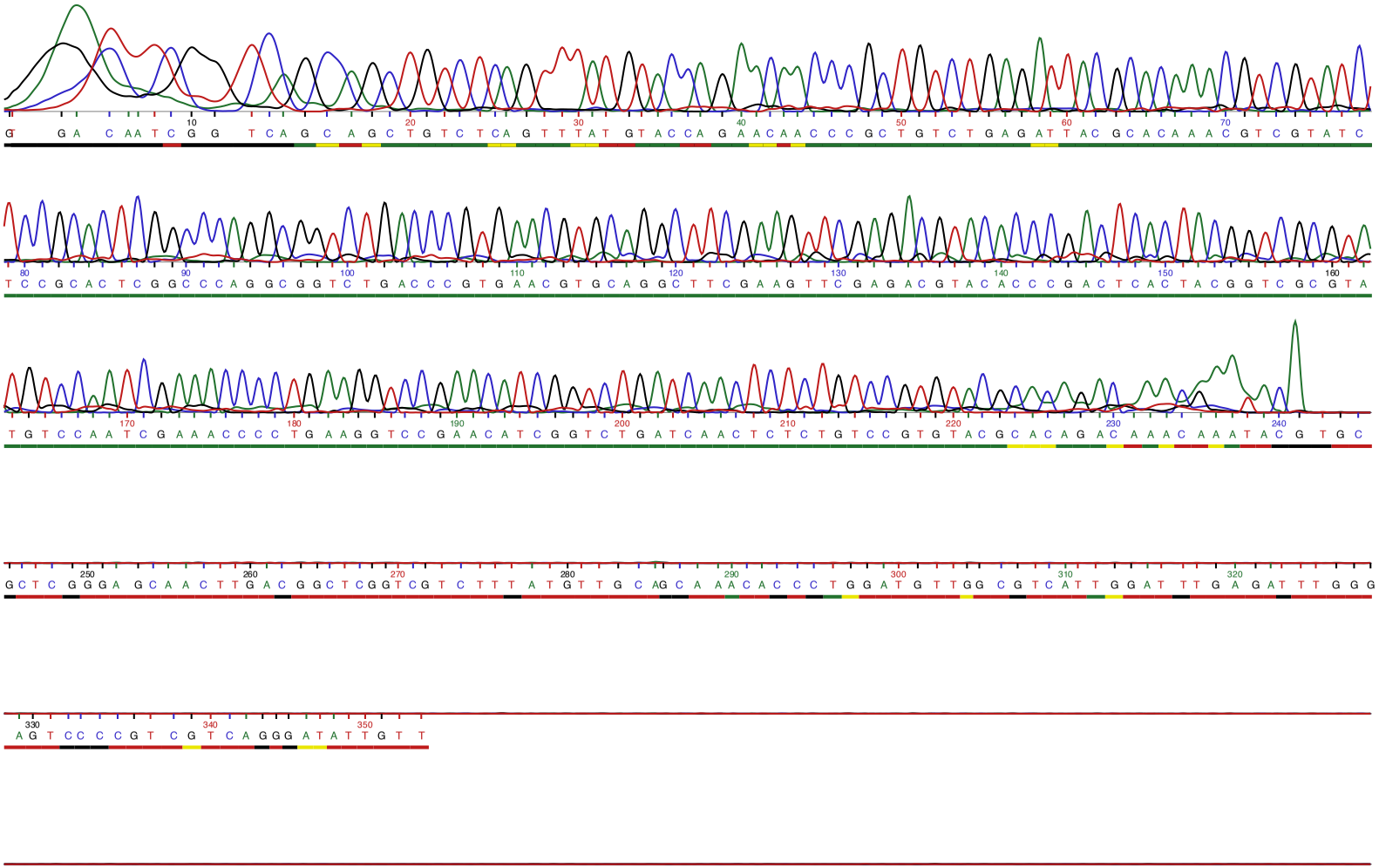

Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182761 to 4182997 GenBank Graphics
Next Match Previous Match
Score 401 bits(217)
Expect 2e-112
Identities 231/237(97%)
Gaps 3/237(1%)
Strand Plus/Plus
Query 8 TCGG-T-CAG-CAGCTGTCTCAGTTTATGTACCAGAACAACCCGCTGTCTGAGATTACGC 64
Sbjct 4182761 TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC 4182820
Query 65 ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG 124
Sbjct 4182821 ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG 4182880
Query 125 AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG 184
Sbjct 4182881 AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG 4182940
Query 185 GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACAAACAAATACG 241
Sbjct 4182941 GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG 4182997

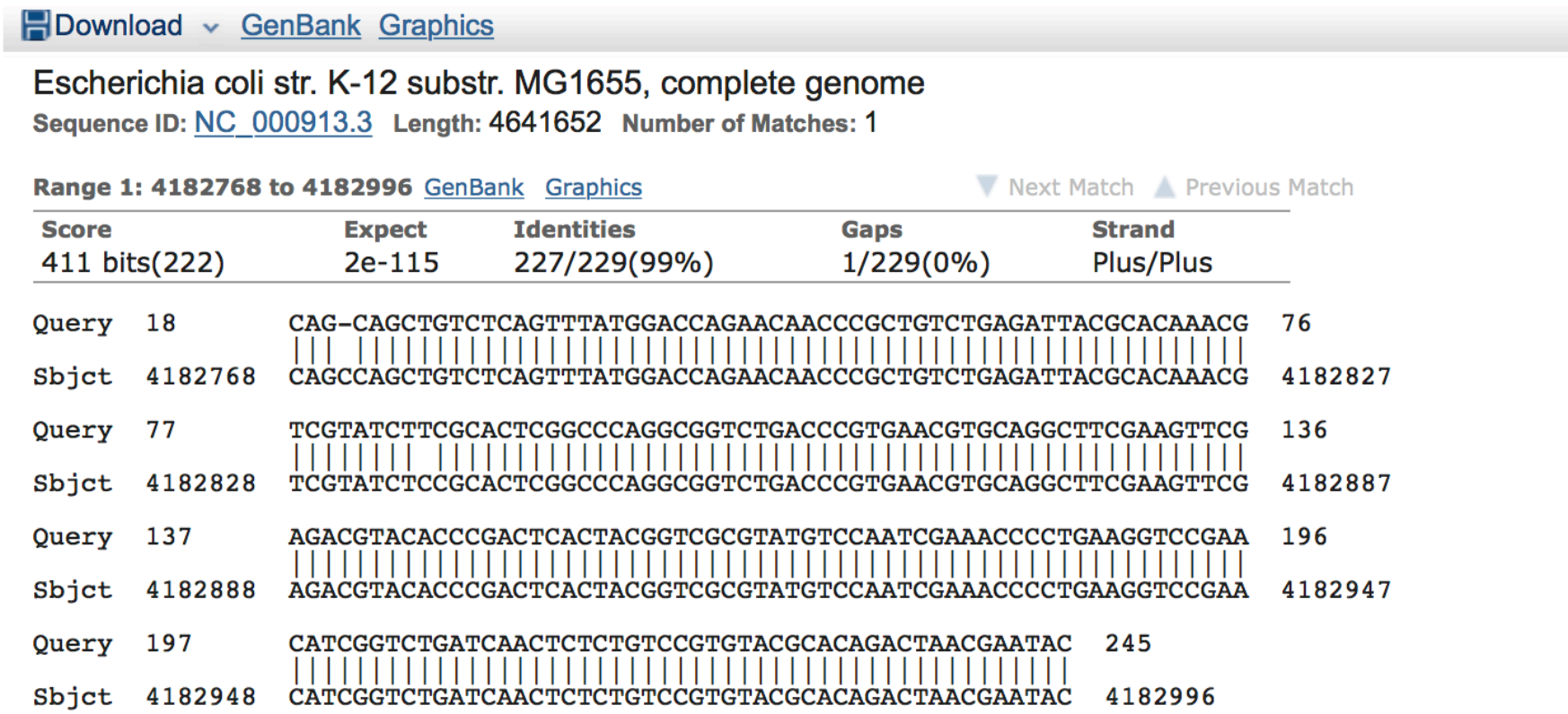
2019/6 CK

Sequence: 50401444

Samples:	16302	Quality:	0 - 9
Bases:	354		10 - 19
Average spacing:	47.0		20 - 29
Average quality >=	10: 98, 20: 23, 30: 200		>= 30

Page: 1 / 3
12.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182996 GenBank Graphics Next Match Previous Match

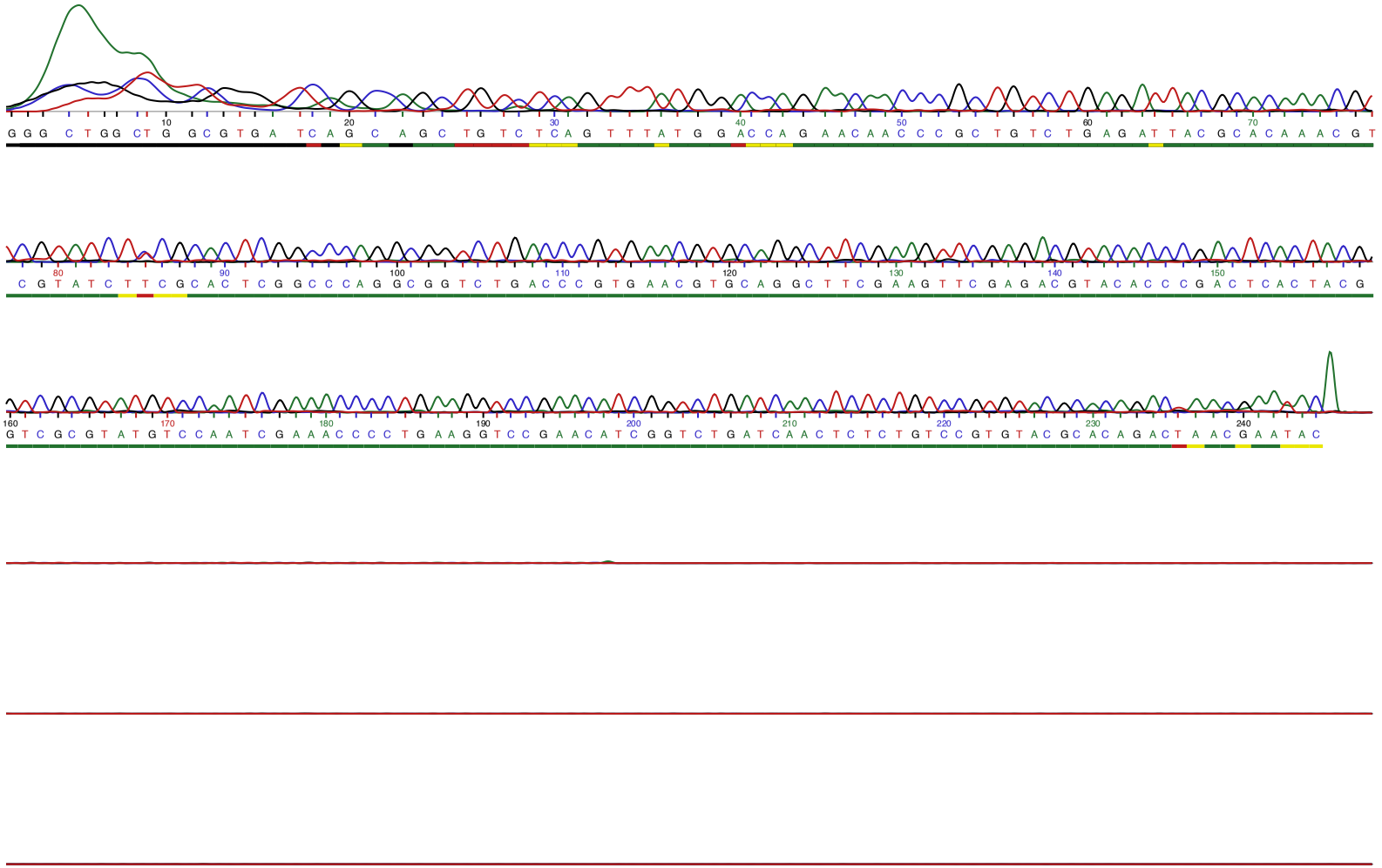
Score	Expect	Identities	Gaps	Strand
411 bits(222)	2e-115	227/229(99%)	1/229(0%)	Plus/Plus

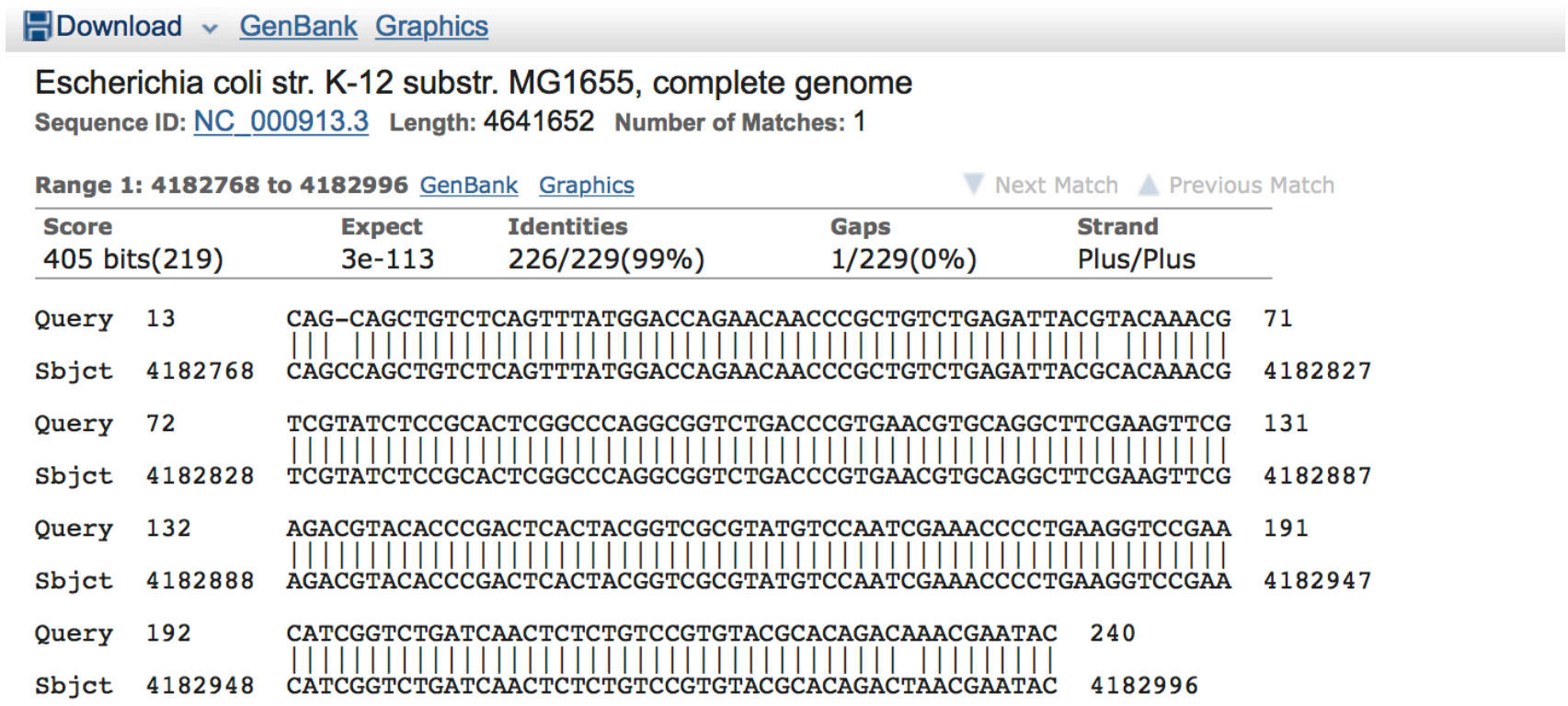
```
Query  18       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  76
                ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  77       TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  136
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  137      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  196
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  197      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  245
                ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/6 CF I

Sequence: 50401451

Samples:	10730	Quality:	0 - 9
Bases:	246		10 - 19
Average spacing:	44.0		20 - 29
Average quality >=	10: 9, 20: 17, 30: 201		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182996 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
405 bits(219)	3e-113	226/229(99%)	1/229(0%)	Plus/Plus

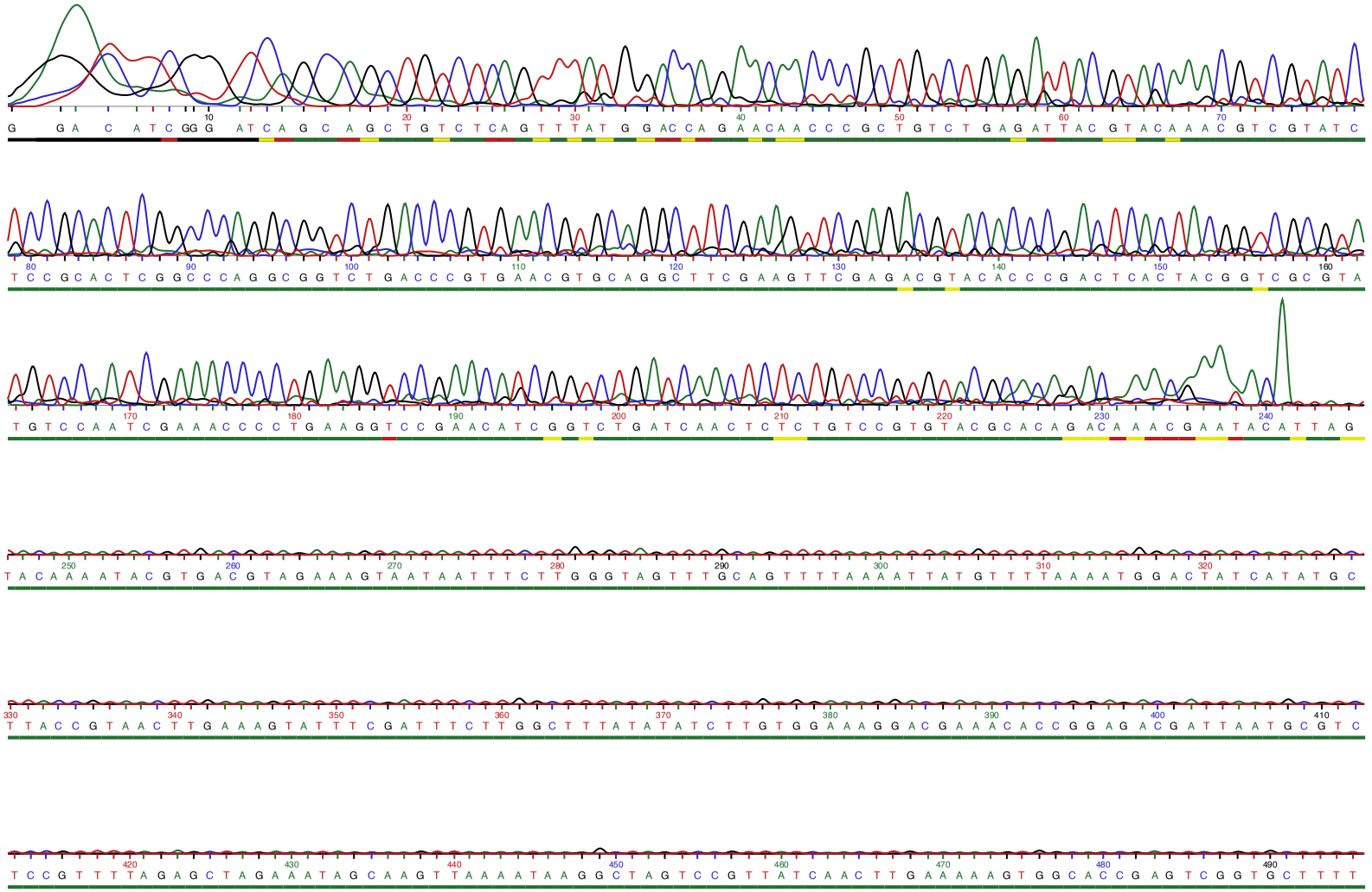
```
Query  13       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGTACAAACG  71
                ||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| 
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  72       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  131
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  132      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  191
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  192      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACAAACGAATAC  240
                |||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/6 CF II

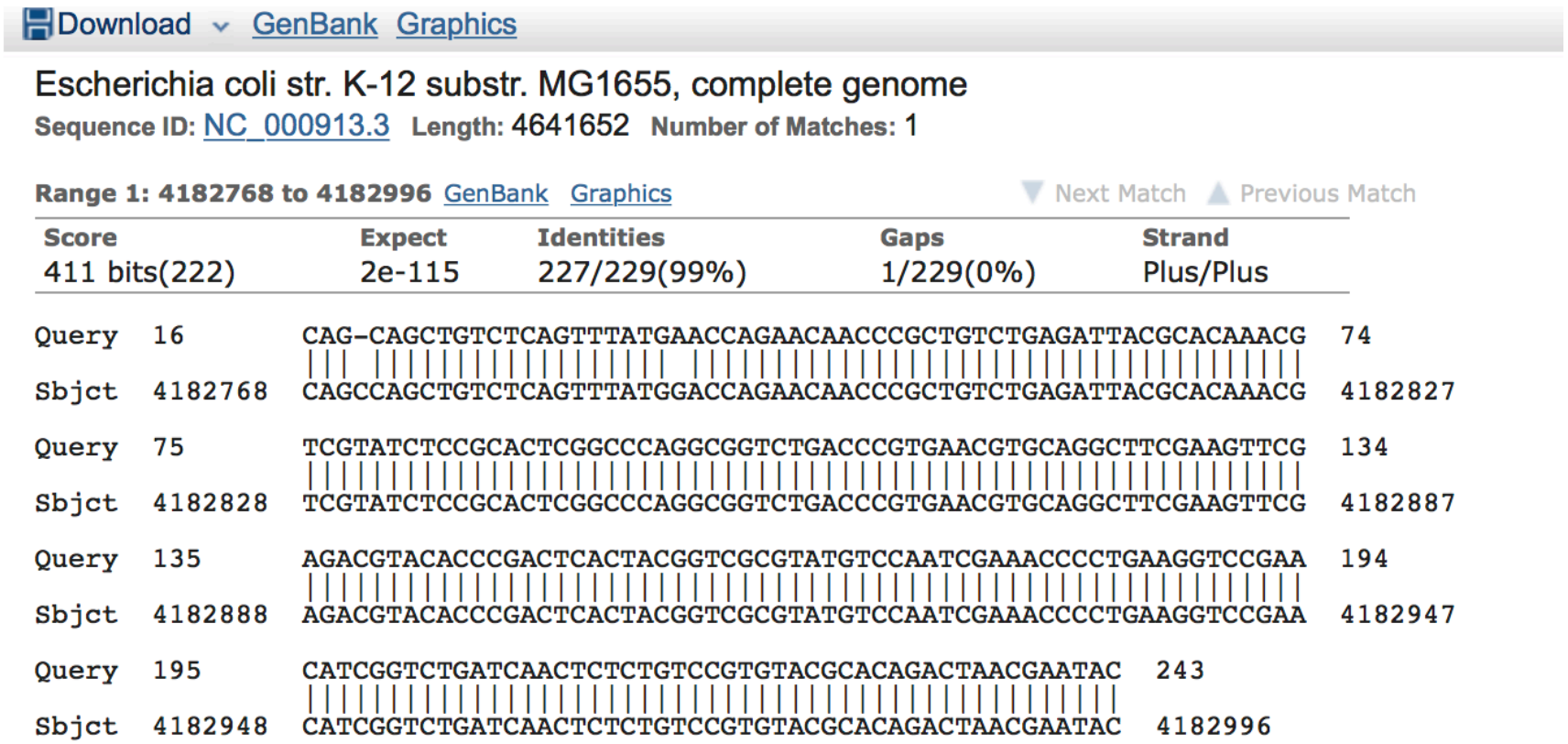
Sequence: 50401468

Samples:	16304	Quality:	0 - 9
Bases:	841		10 - 19
Average spacing:	20.0		20 - 29
Average quality >=	10: 65, 20: 93, 30: 658		>= 30

12.03.2019



2019/6 DK

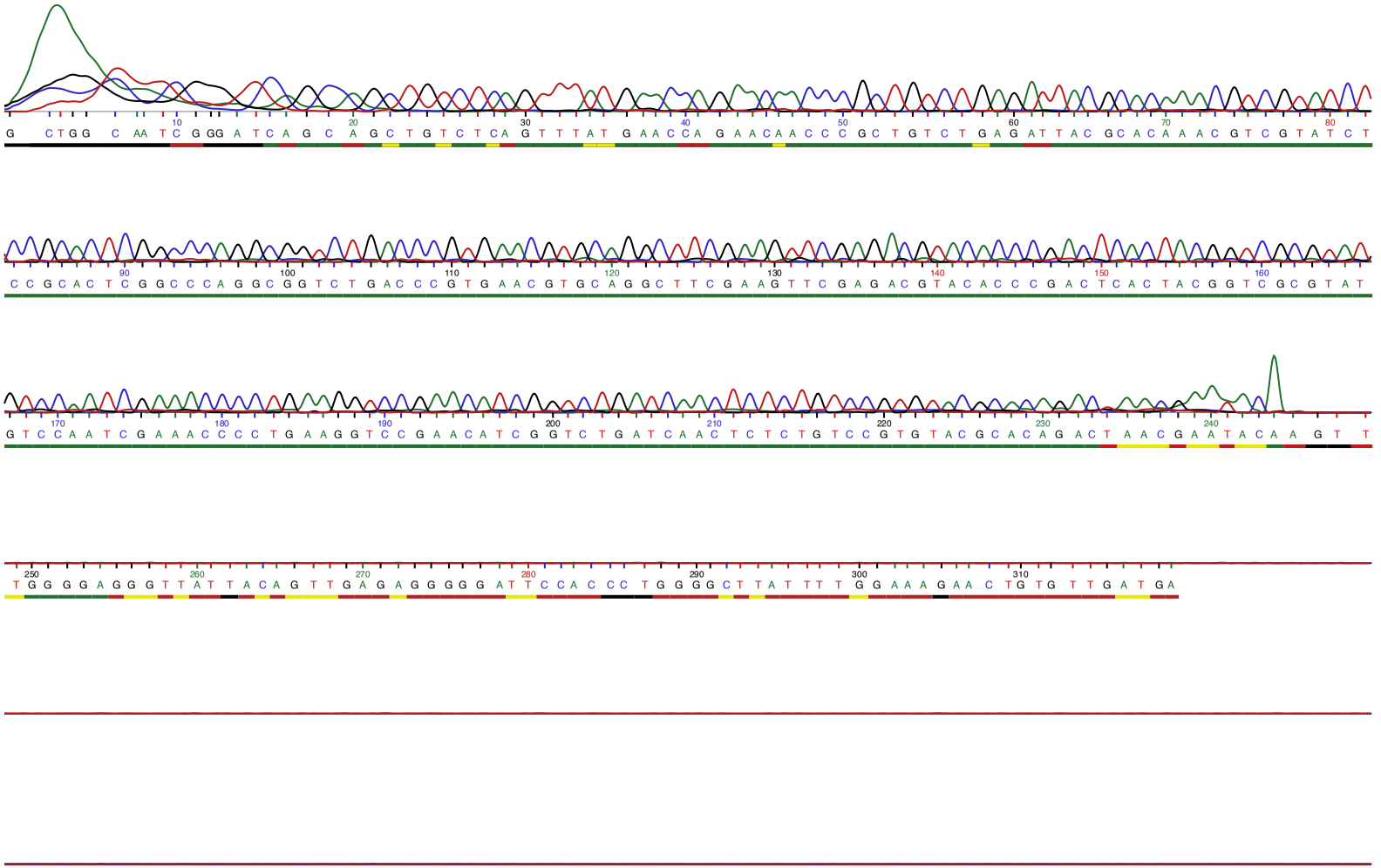

Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182768 to 4182996 GenBank Graphics
Next Match Previous Match
Score 411 bits(222)
Expect 2e-115
Identities 227/229(99%)
Gaps 1/229(0%)
Strand Plus/Plus
Query 16 CAG-CAGCTGTCTCAGTTTATGAACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 74
Sbjct 4182768 CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 4182827
Query 75 TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 134
Sbjct 4182828 TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 4182887
Query 135 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 194
Sbjct 4182888 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 4182947
Query 195 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC 243
Sbjct 4182948 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC 4182996

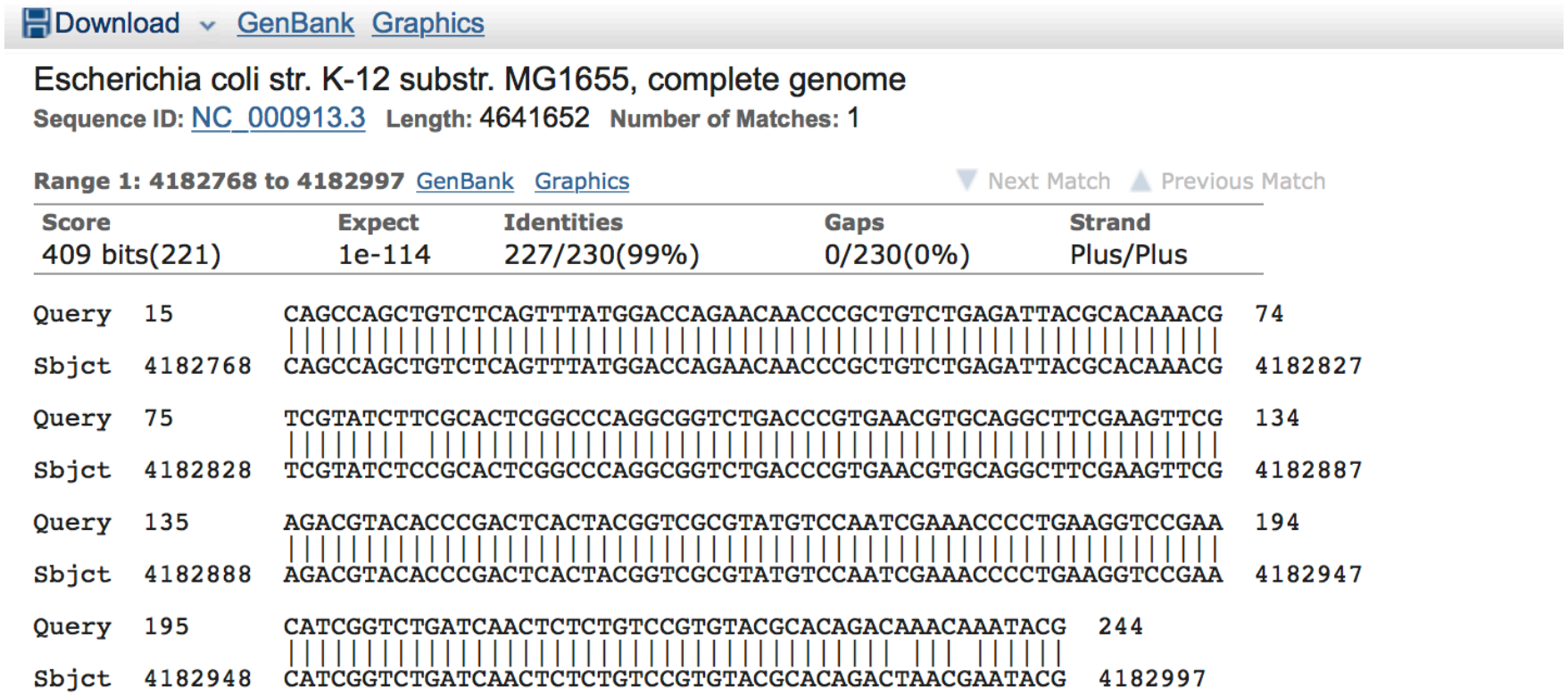
2019/6 DK

Sequence: 50401475

Samples:	16300	Quality:	0 - 9
Bases:	320		10 - 19
Average spacing:	51.0		20 - 29
Average quality >=	10: 60, 20: 30, 30: 210		>= 30

Page: 1 / 3
12.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
409 bits(221)	1e-114	227/230(99%)	0/230(0%)	Plus/Plus

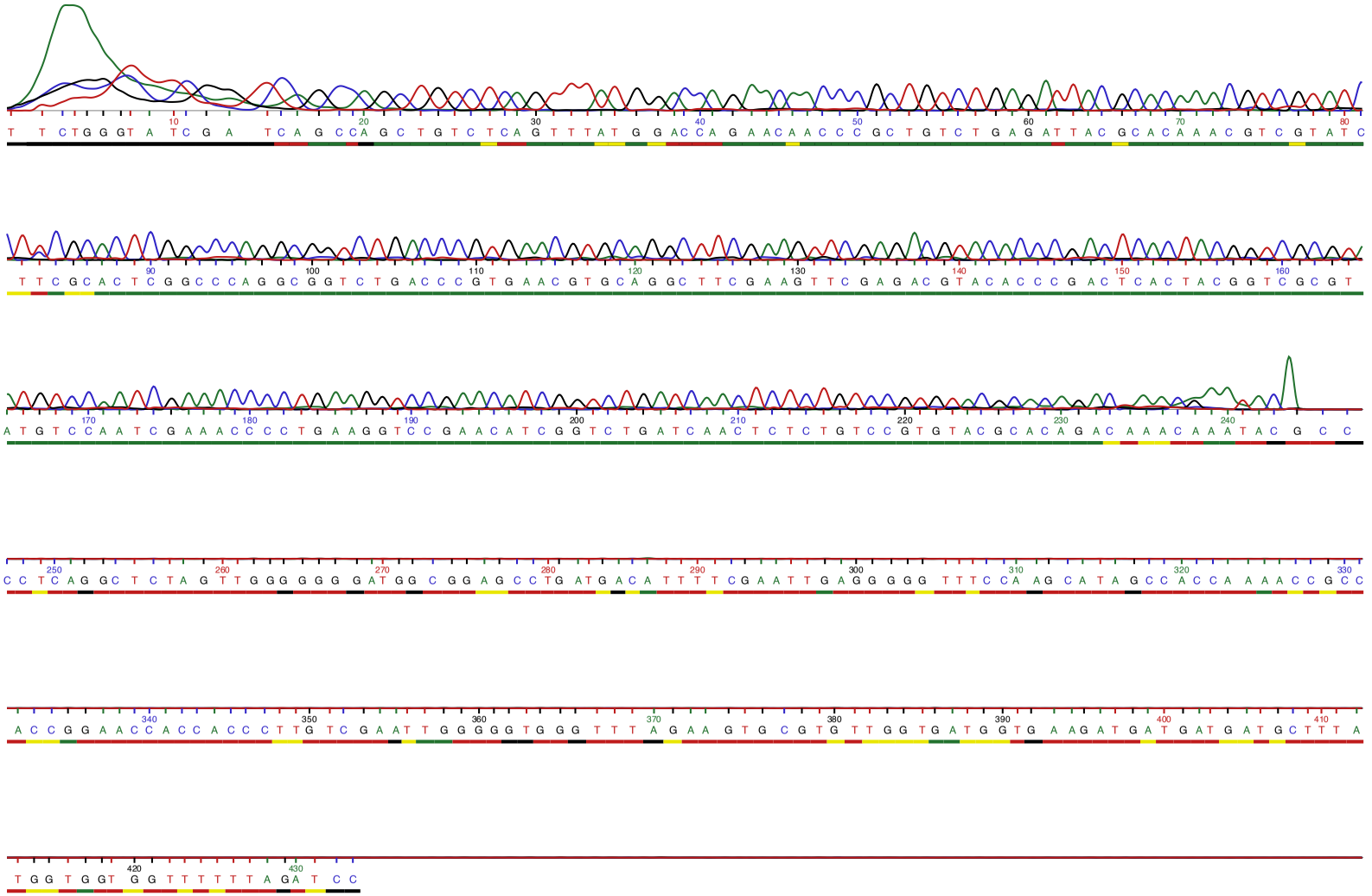
```
Query  15       CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  74
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  134
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  135      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  195      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACAAACAAATACG  244
                |||||||||||||||||||||||||||||||||||| ||| ||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

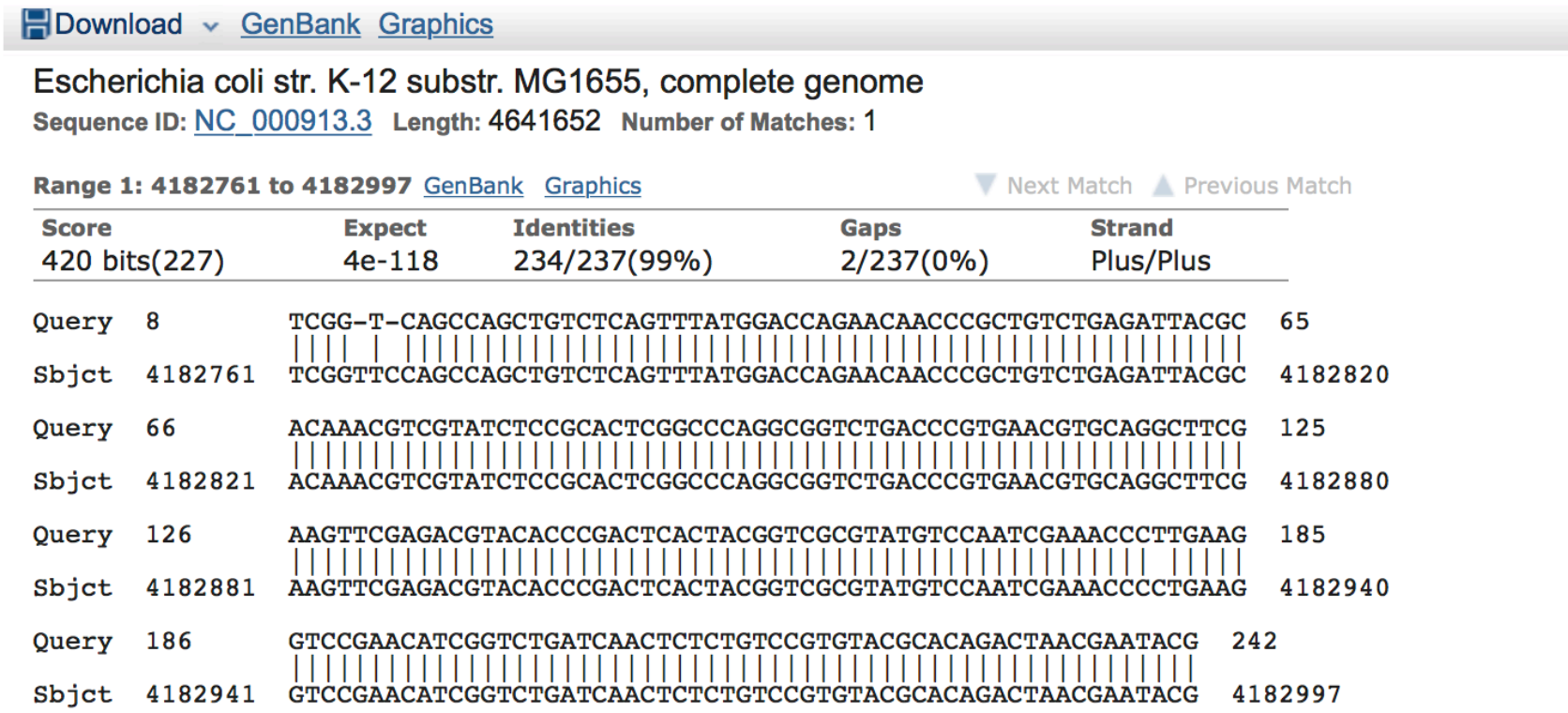
2019/6 DF I

Sequence: 50401482

Samples:	16301	Quality:	0 - 9
Bases:	434		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 146, 20: 47, 30: 207		>= 30

12.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
420 bits(227)	4e-118	234/237(99%)	2/237(0%)	Plus/Plus

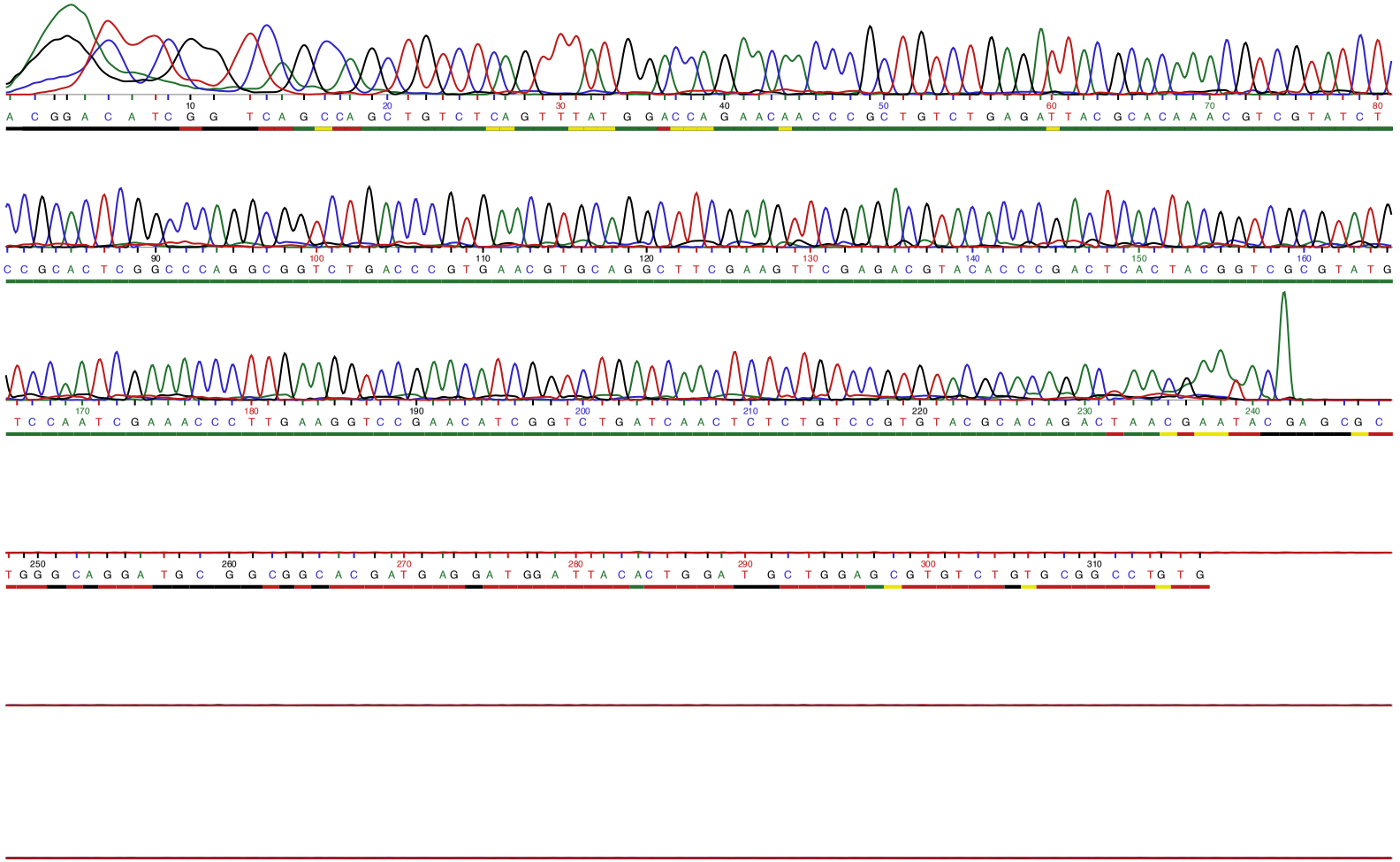
```
Query  8        TCGG-T-CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  65
               |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  66       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  126      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCTTGAAG  185
               |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  186      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  242
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

2019/6 DF II

Sequence: 07456794

Samples:	16302	Quality: 0 - 9	
Bases:	317	10 - 19	
Average spacing:	52.0	20 - 29	
Average quality >=	10: 62, 20: 18, 30: 207	>= 30	

12.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
409 bits(221)	1e-114	232/237(98%)	2/237(0%)	Plus/Plus

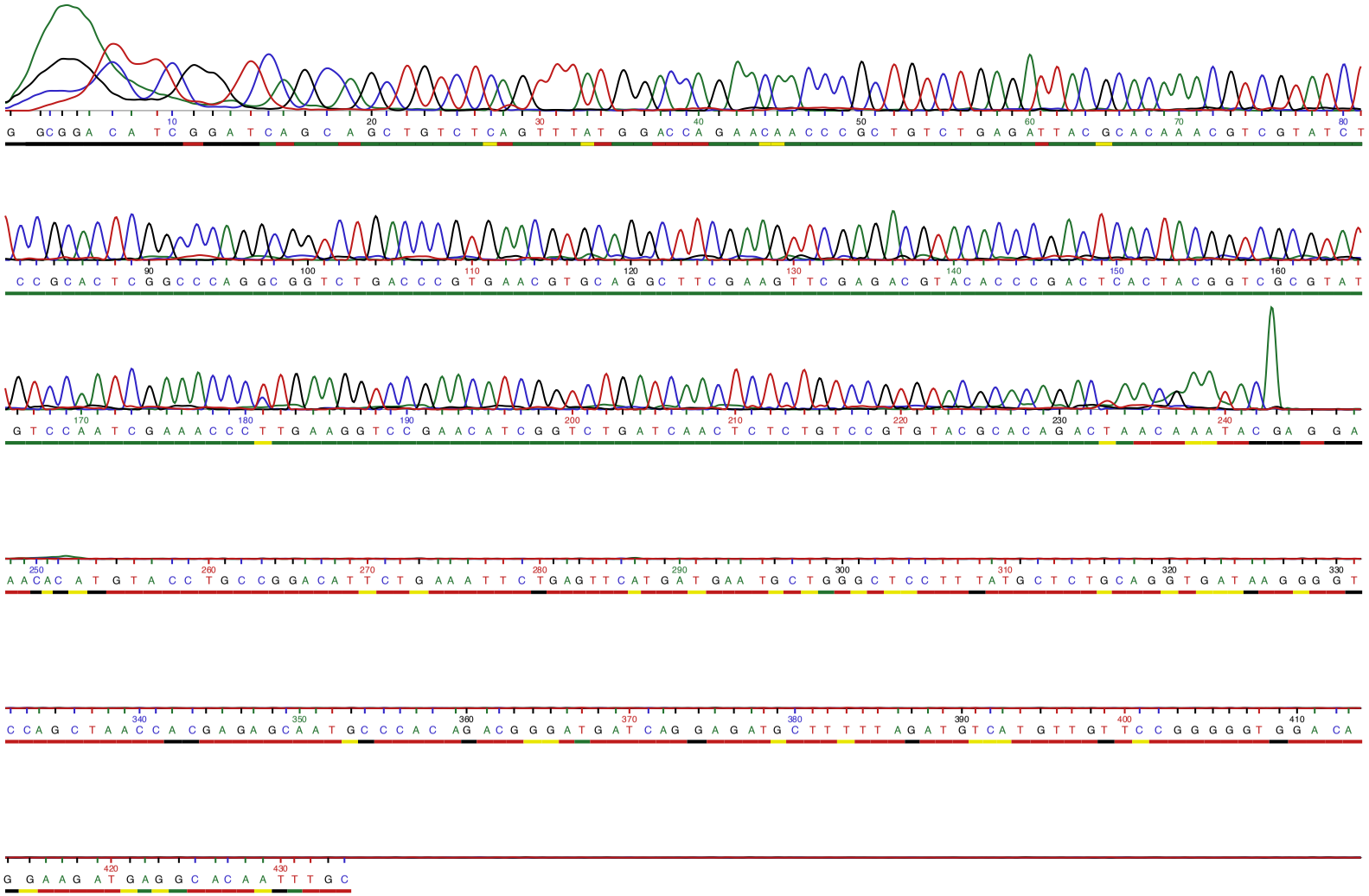
```
Query  9        TCGGAT-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  66
               |||| | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  67       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  127      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCTTGAAG  186
               |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  187      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATACG  243
               ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

2019/8 AF I

Sequence: 10312537

Samples:	16303	Quality:	0 - 9
Bases:	435		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 151, 20: 39, 30: 209		>= 30

Page: 1 / 3
12.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182996 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
424 bits(229)	4e-119	234/236(99%)	1/236(0%)	Plus/Plus

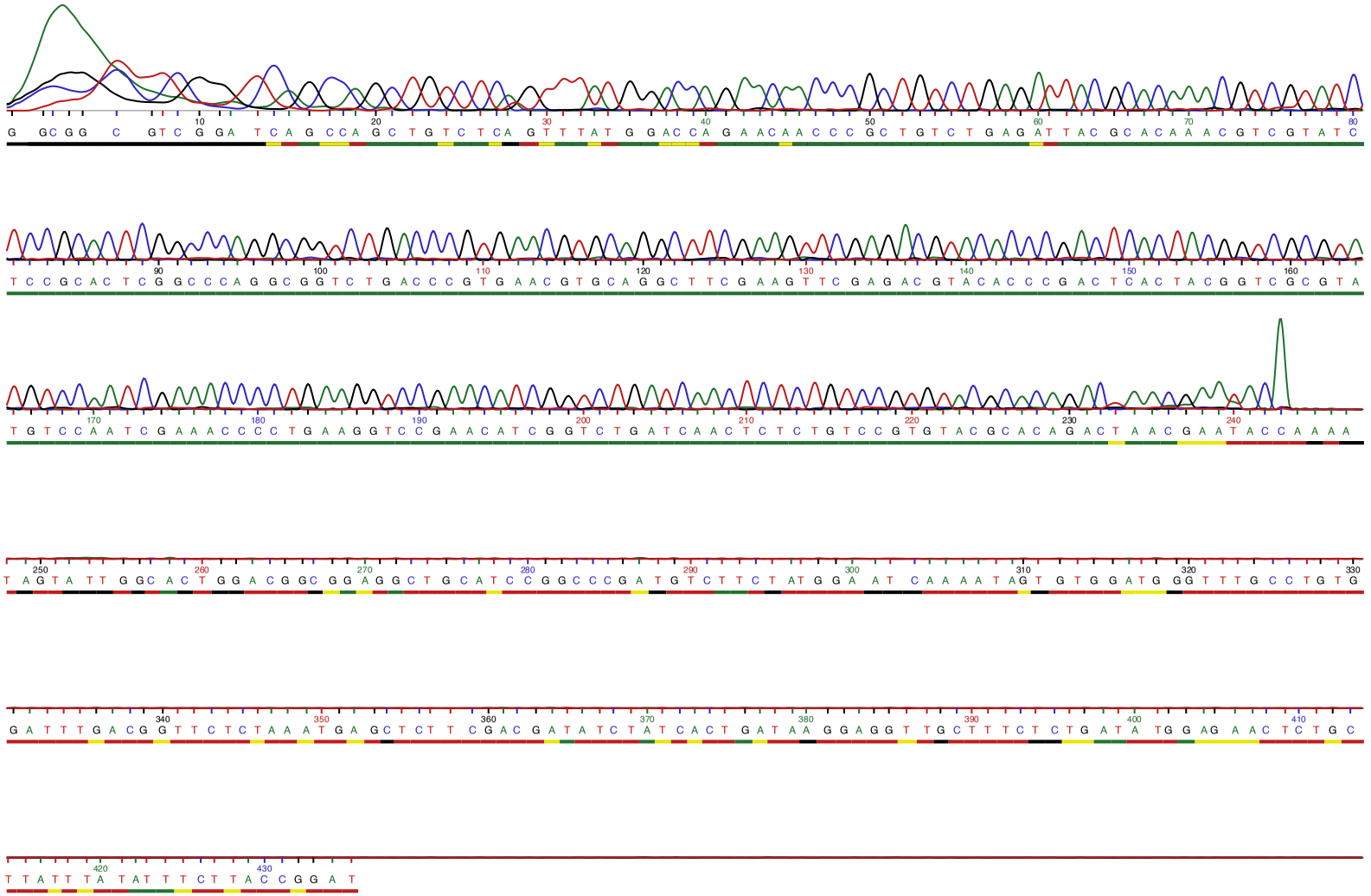
```
Query  8        TCGGAT-CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  66
               |||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  67       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  127      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  187      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  242
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/8 BF I

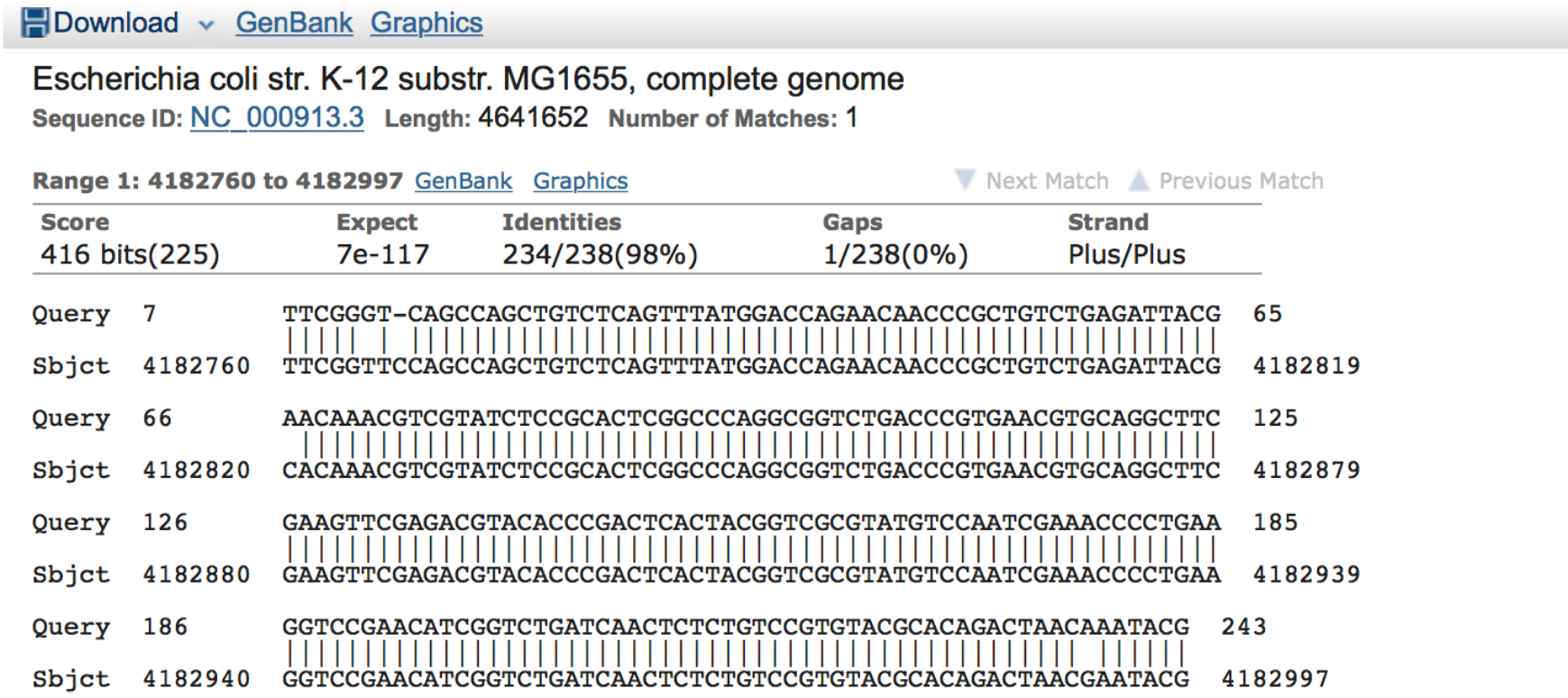
Sequence: 12198825

Samples:	16302	Quality:	0 - 9
Bases:	436		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 135, 20: 47, 30: 217		>= 30

12.03.2019



2019/8 CK

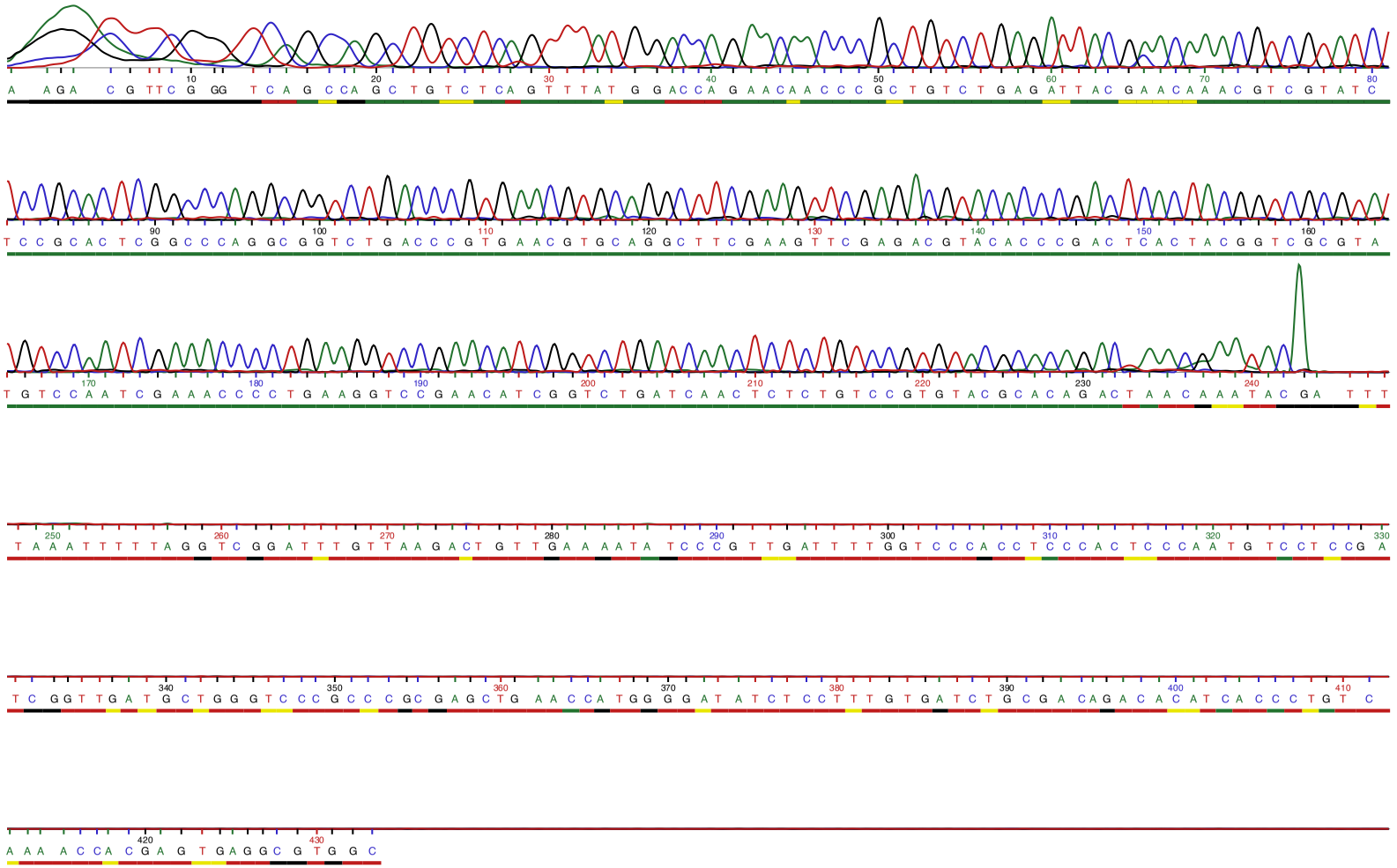

Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182760 to 4182997 GenBank Graphics
Next Match Previous Match
Score 416 bits(225)
Expect 7e-117
Identities 234/238(98%)
Gaps 1/238(0%)
Strand Plus/Plus
Query 7 TTCGGGT-CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG 65
Sbjct 4182760 TTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG 4182819
Query 66 AACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC 125
Sbjct 4182820 CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC 4182879
Query 126 GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA 185
Sbjct 4182880 GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA 4182939
Query 186 GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATACG 243
Sbjct 4182940 GGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG 4182997

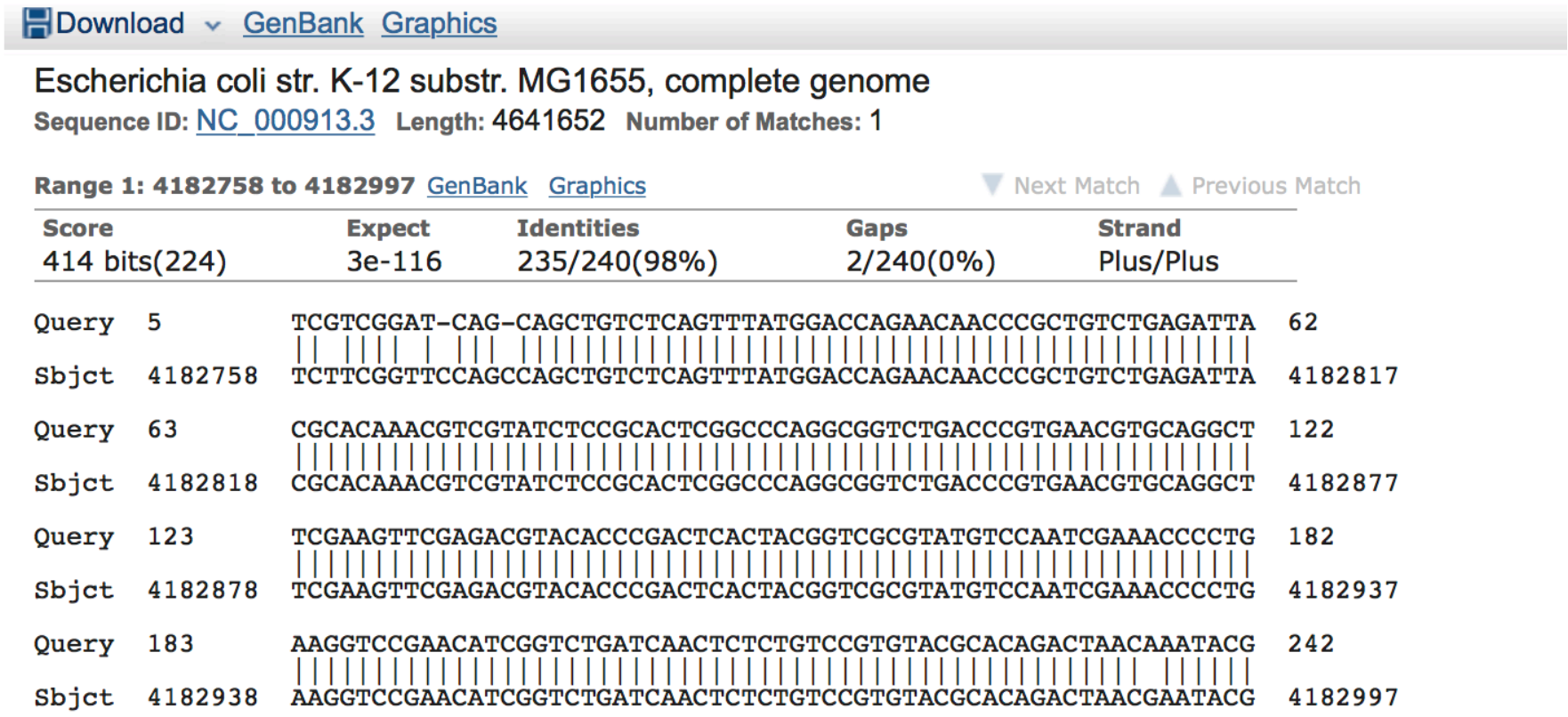
2019/8 CK

Sequence: 13990442

Samples:	16304	Quality:	0 - 9
Bases:	434		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 152, 20: 40, 30: 205		>= 30

12.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182758 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
414 bits(224)	3e-116	235/240(98%)	2/240(0%)	Plus/Plus

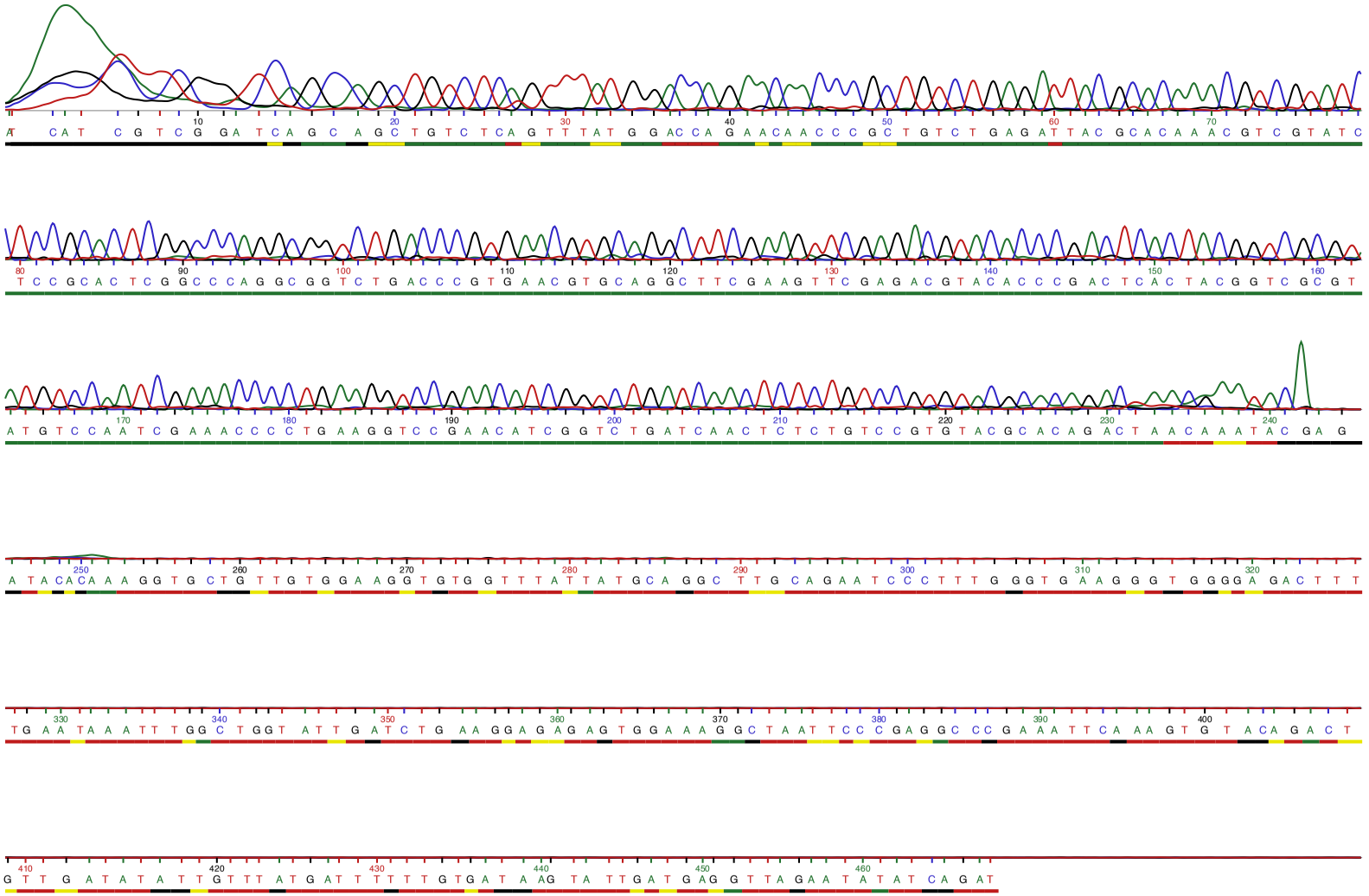
```
Query  5        TCGTCGGAT-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTA  62
               ||  ||||  | |||  ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182758  TCTTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTA  4182817

Query  63       CGCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCT  122
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182818  CGCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCT  4182877

Query  123      TCGAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTG  182
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182878  TCGAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTG  4182937

Query  183      AAGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATACG  242
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  4182938  AAGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

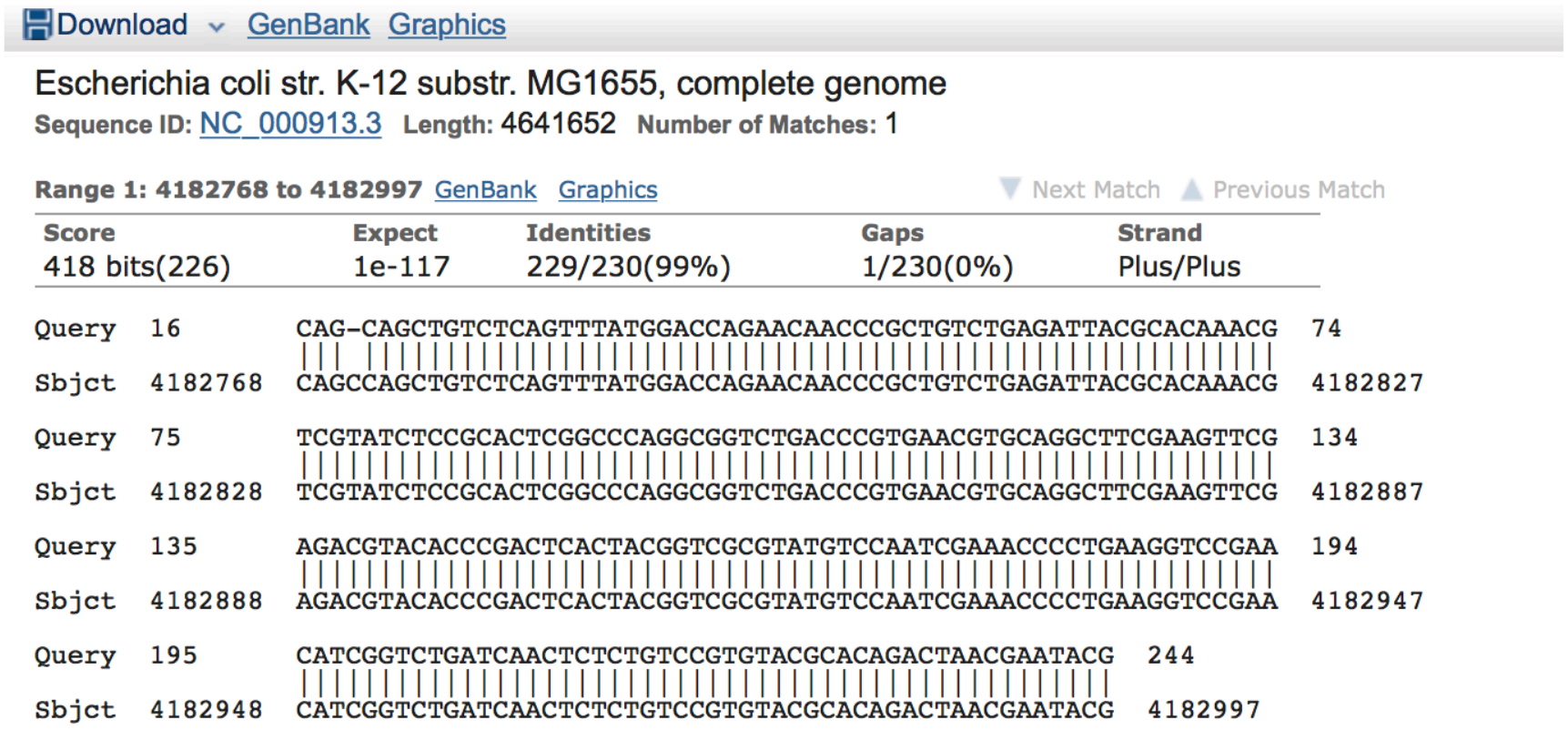
2019/8 CF I

Sequence: 14285165

Samples:	16303	Quality:	0 - 9
Bases:	469		10 - 19
Average spacing:	35.0		20 - 29
Average quality >=	10: 168, 20: 44, 30: 211		>= 30

12.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182997 GenBank Graphics Next Match Previous Match

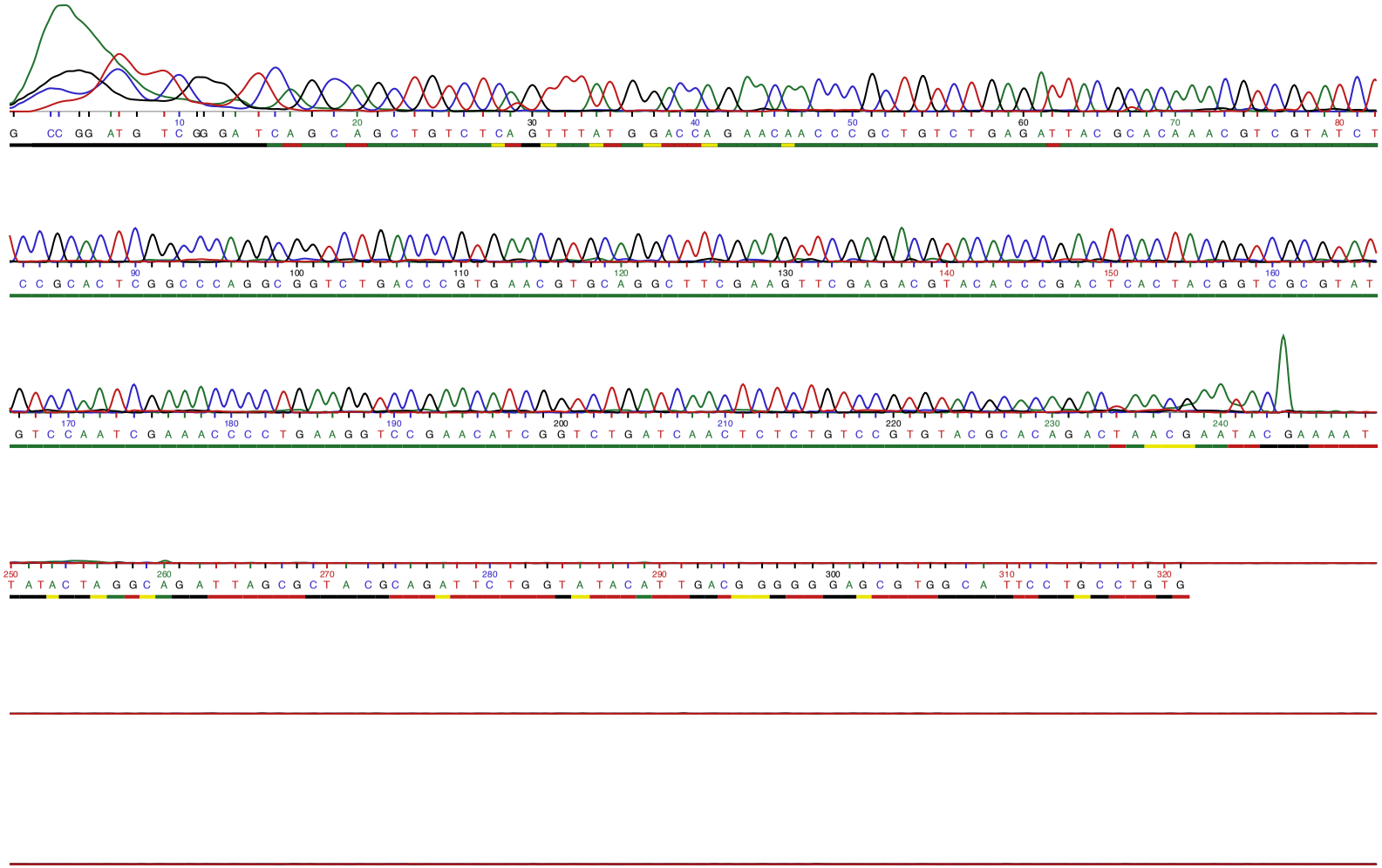
Score	Expect	Identities	Gaps	Strand
418 bits(226)	1e-117	229/230(99%)	1/230(0%)	Plus/Plus

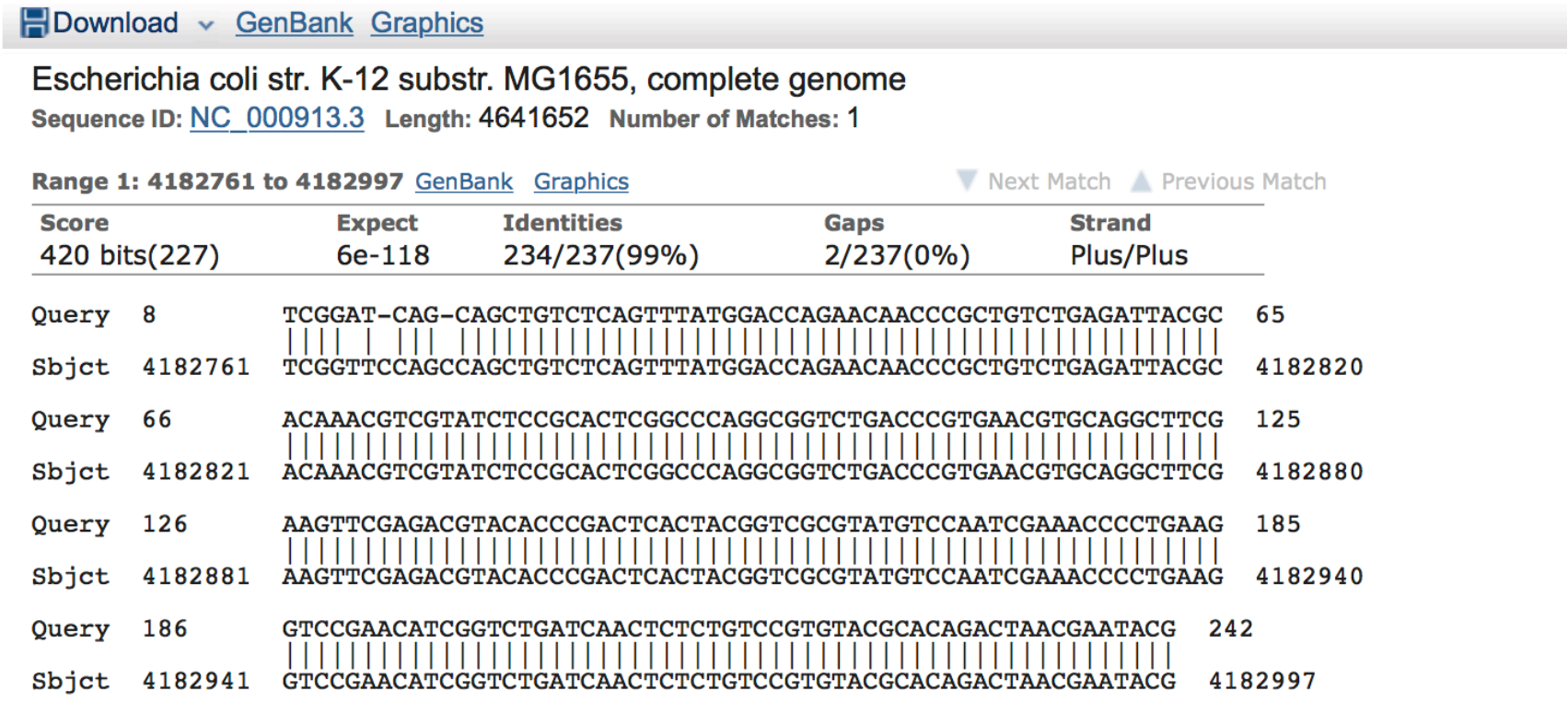
```
Query  16       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  74
                ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  135      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  195      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  244
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

2019/8 CF II

Sequence: 16675674

Samples:	16302	Quality:	0 - 9
Bases:	322		10 - 19
Average spacing:	51.0		20 - 29
Average quality >=	10: 49, 20: 18, 30: 209		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182997 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
420 bits(227)	6e-118	234/237(99%)	2/237(0%)	Plus/Plus

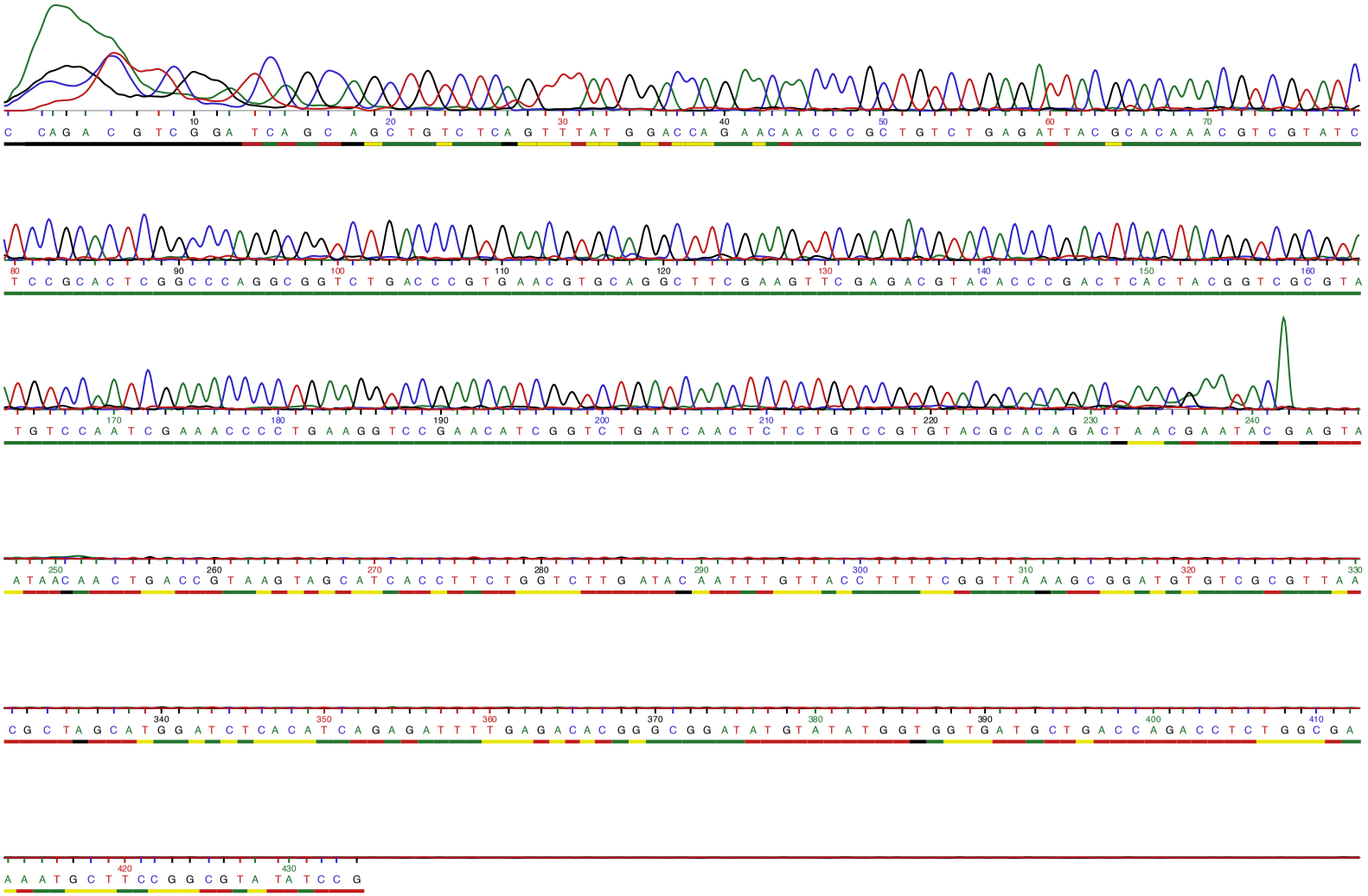
```
Query  8        TCGGAT-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  65
               |||| | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  66       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  126      AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  185
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  186      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  242
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATACG  4182997
```

2019/8 DK

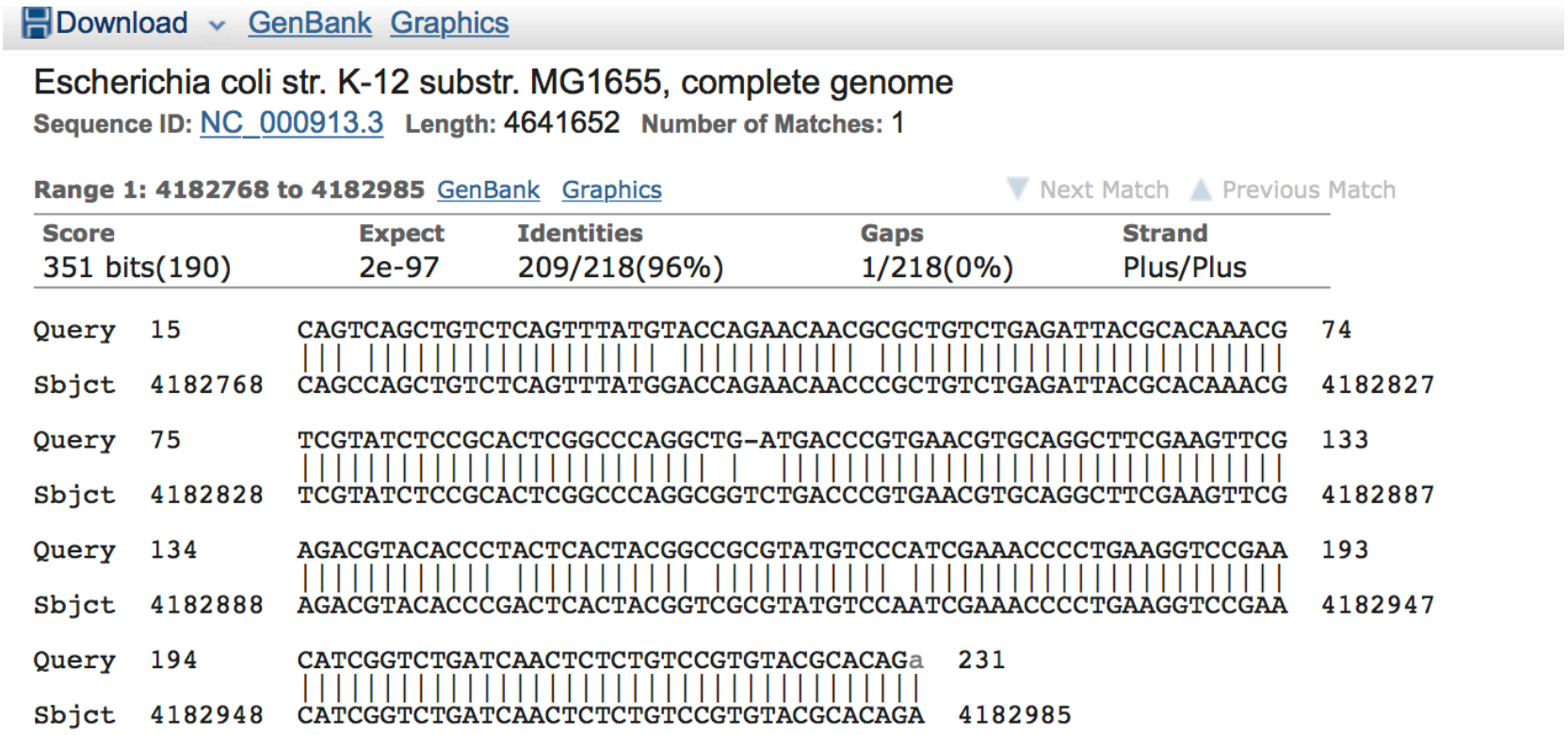
Sequence: 18524628

Samples:	16300	Quality:	0 - 9
Bases:	435		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 92, 20: 69, 30: 252		>= 30

Page: 1 / 3
12.03.2019



2019/10 AK


Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182768 to 4182985 GenBank Graphics
Next Match Previous Match
Score 351 bits(190)
Expect 2e-97
Identities 209/218(96%)
Gaps 1/218(0%)
Strand Plus/Plus
Query 15 CAGTCAGCTGTCTCAGTTTATGTACCAGAACAACGCGCTGTCTGAGATTACGCACAAACG 74
Sbjct 4182768 CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 4182827
Query 75 TCGTATCTCCGCACTCGGCCCAGGCTG-ATGACCCGTGAACGTGCAGGCTTCGAAGTTCG 133
Sbjct 4182828 TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 4182887
Query 134 AGACGTACACCCTACTCACTACGGCCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAA 193
Sbjct 4182888 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 4182947
Query 194 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGa 231
Sbjct 4182948 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA 4182985

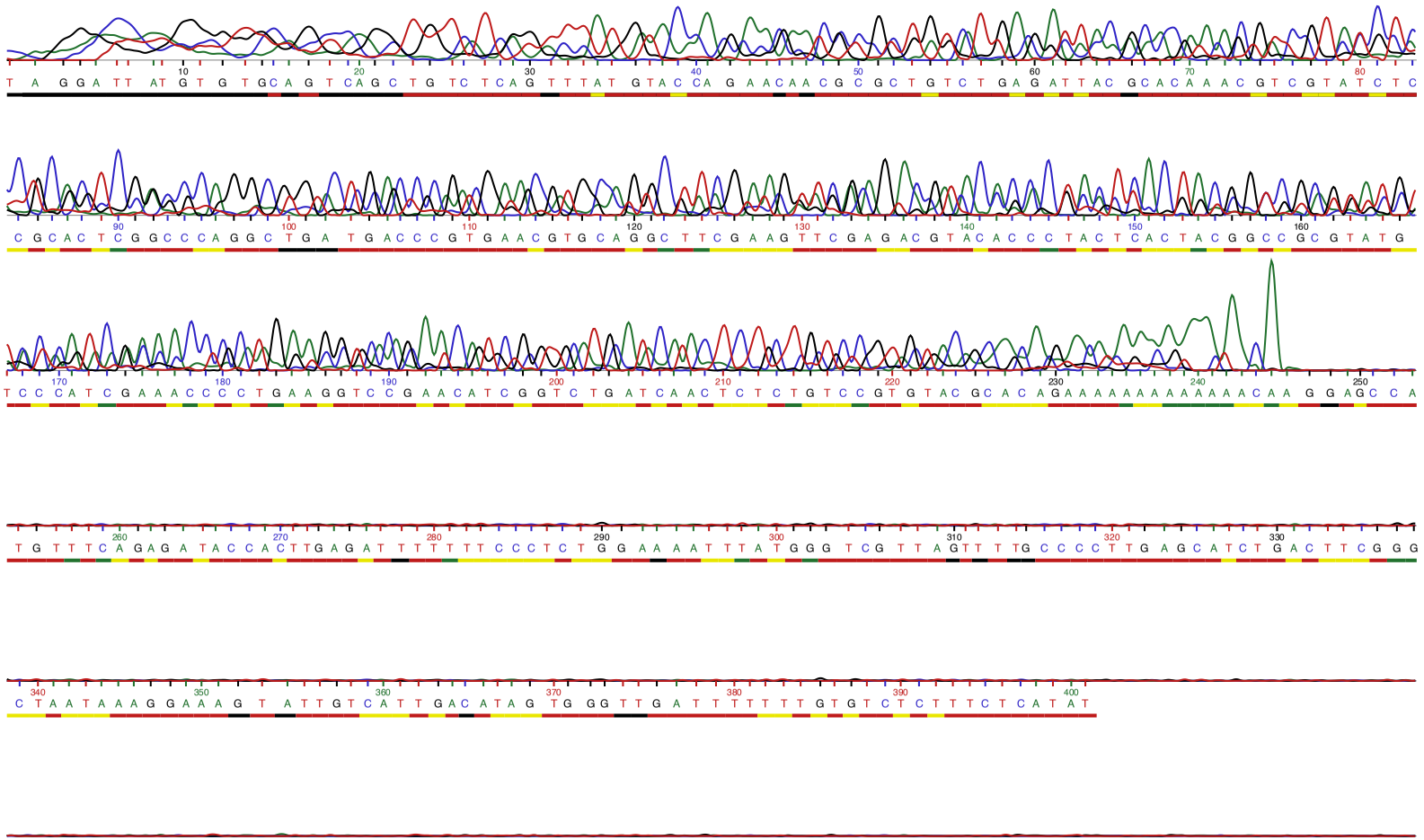
Sequence: 50401512

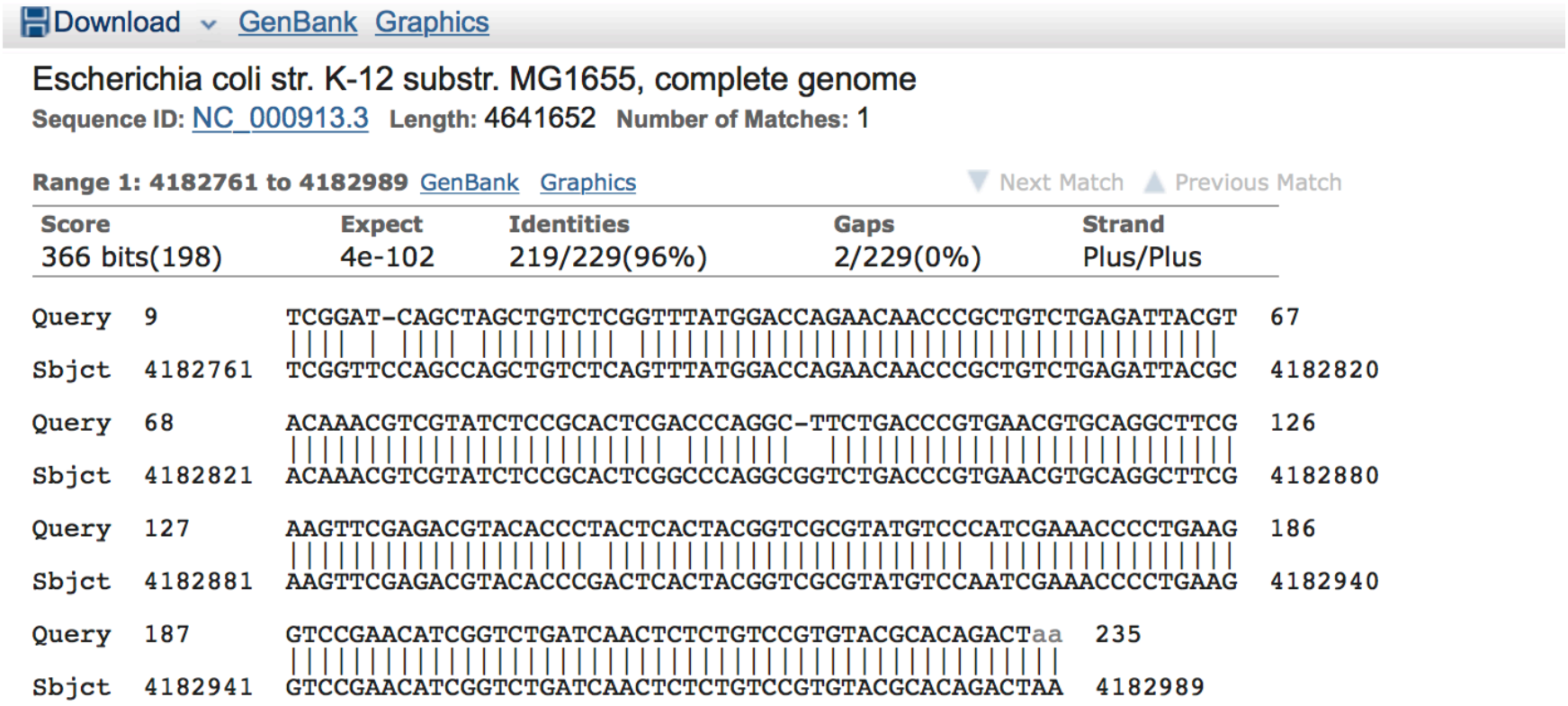
2019/10 AK

Samples: 16302
Bases: 402
Average spacing: 41.0
Average quality >= 10: 234, 20: 105, 30: 24

Quality: 0 - 9
10 - 19
20 - 29
>= 30

01.04.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182989 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
366 bits(198)	4e-102	219/229(96%)	2/229(0%)	Plus/Plus

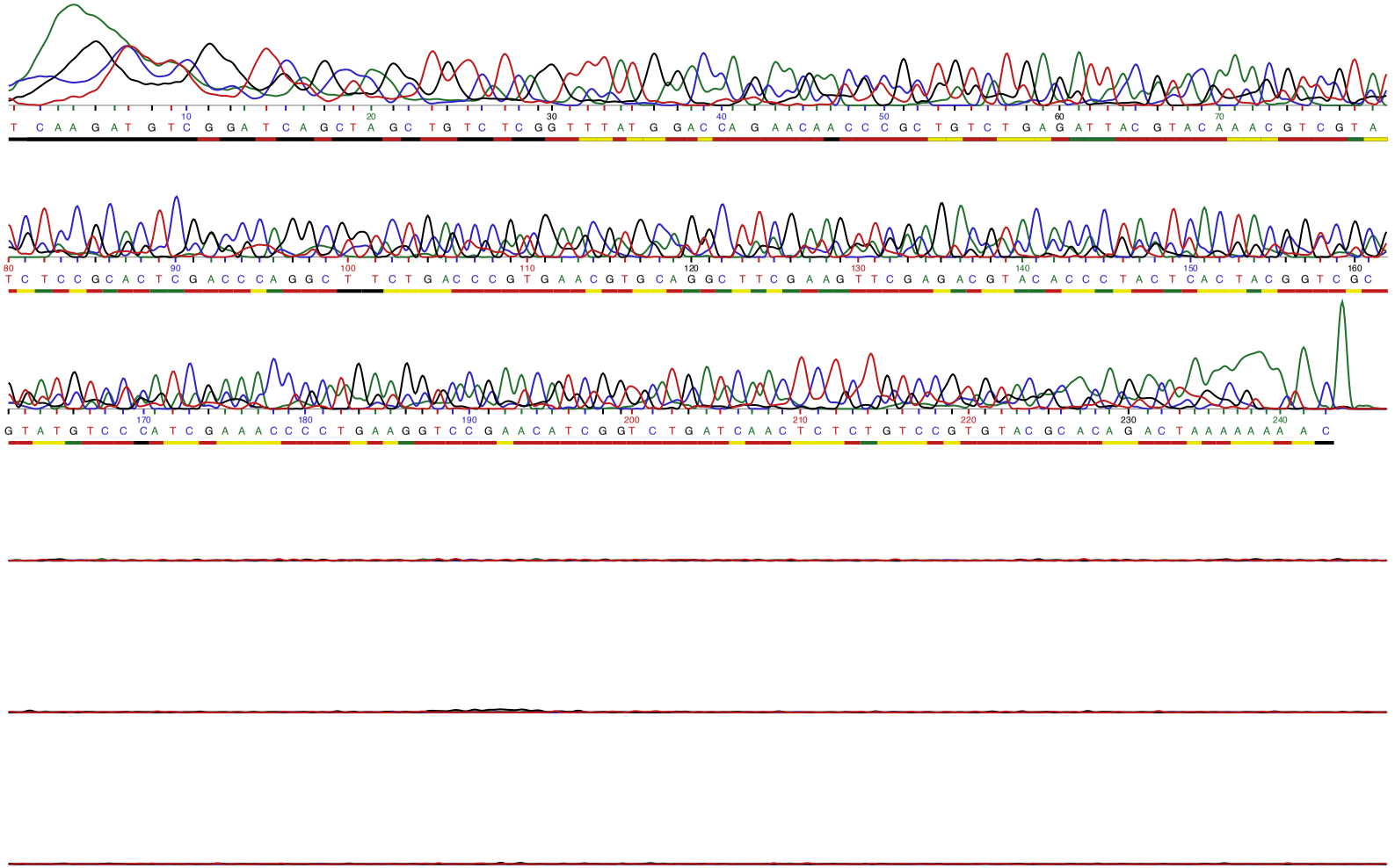
```
Query  9        TCGGAT-CAGCTAGCTGTCTCGGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGT  67
               |||| | |||| |||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  68       ACAAACGTCGTATCTCCGCACTCGACCCAGGC-TTCTGACCCGTGAACGTGCAGGCTTCG  126
               |||||||||||||||||||||||| |||||||  |||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  127      AAGTTCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAG  186
               |||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  187      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  235
               ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

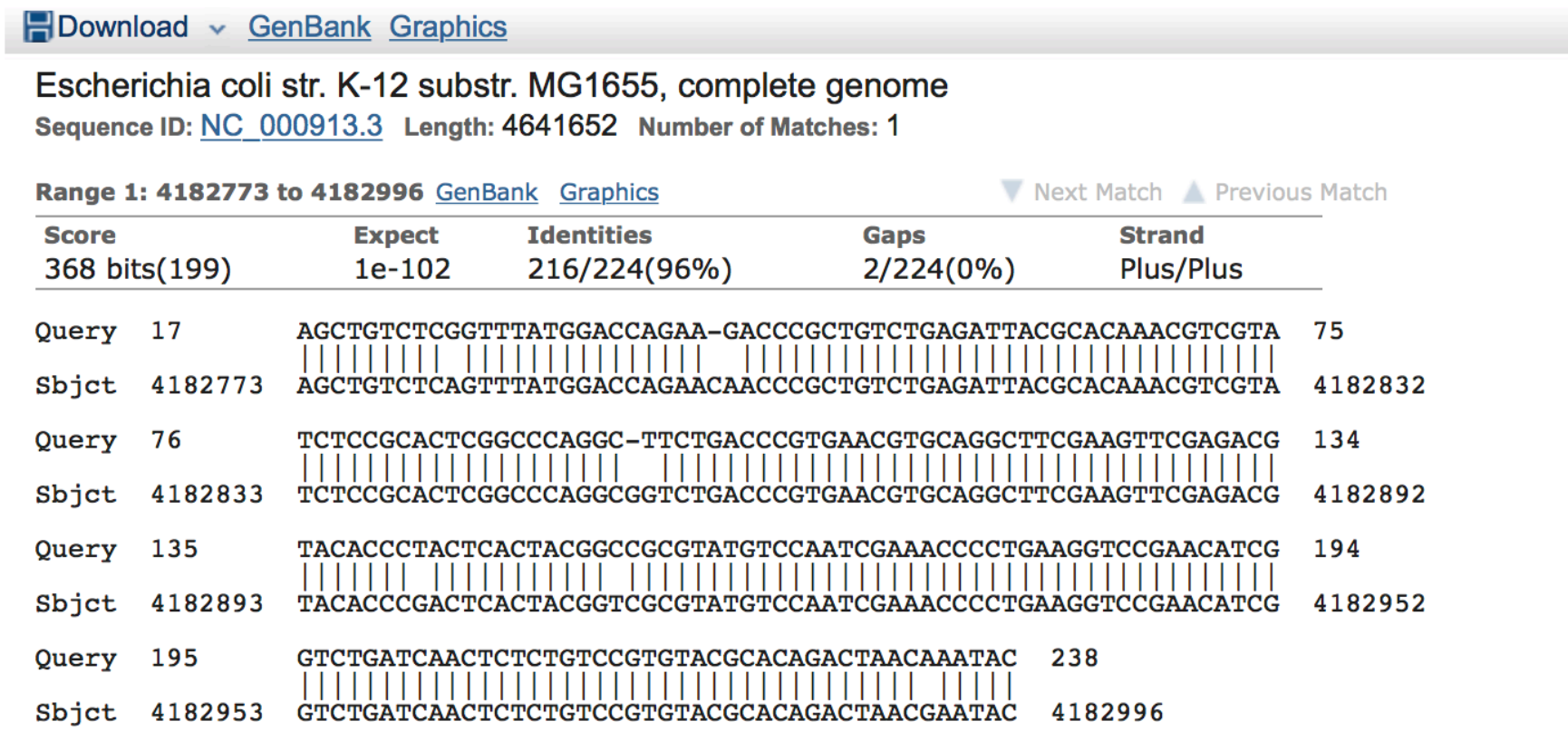
2019/10 AF I

Sequence: 50401505

Samples:	15441	Quality:	0 - 9
Bases:	243		10 - 19
Average spacing:	64.0		20 - 29
Average quality >=	10: 128, 20: 63, 30: 24		>= 30

29.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182773 to 4182996 GenBank Graphics Next Match Previous Match

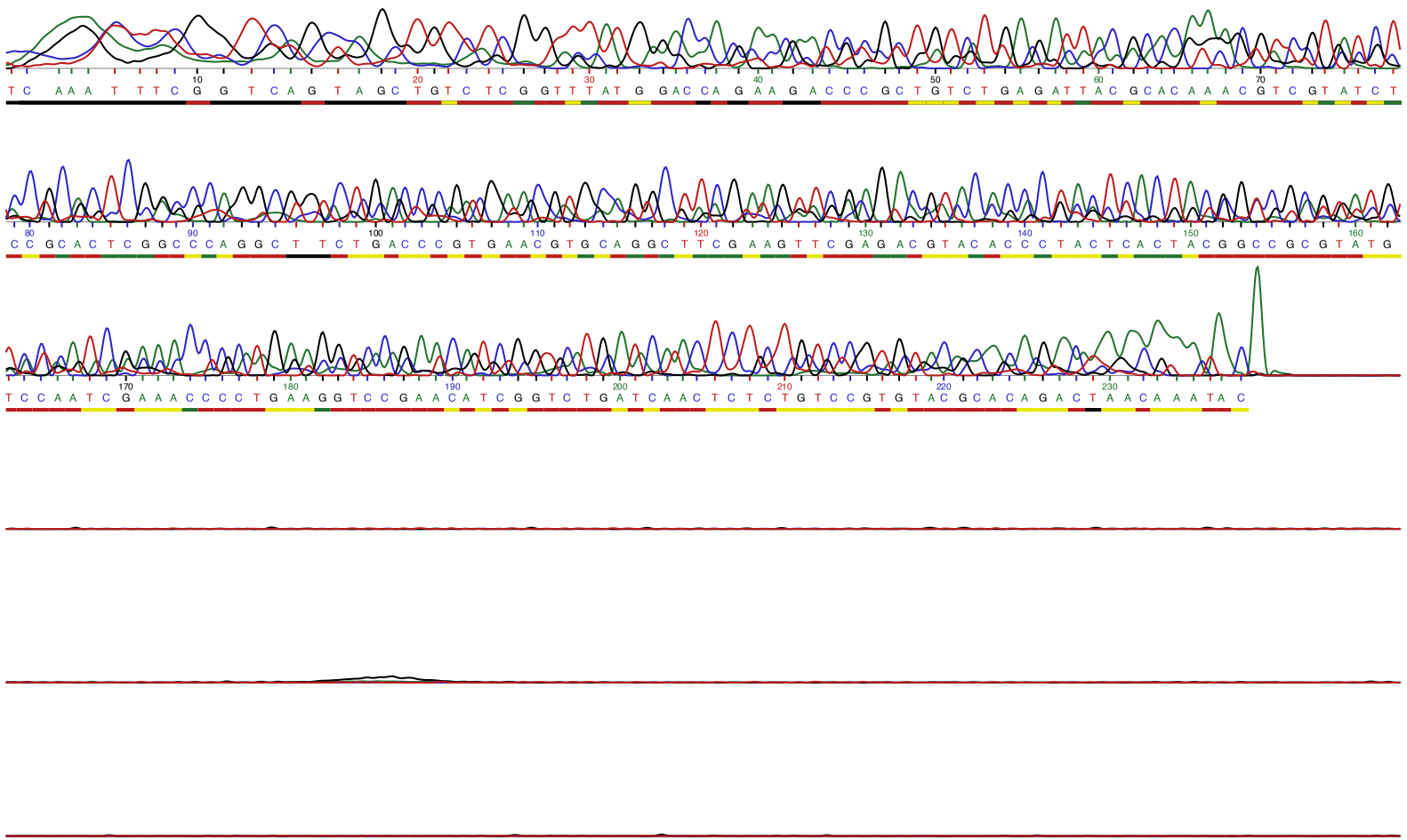
Score	Expect	Identities	Gaps	Strand
368 bits(199)	1e-102	216/224(96%)	2/224(0%)	Plus/Plus

```
Query  17       AGCTGTCTCGGTTTATGGACCAGAA-GACCCGCTGTCTGAGATTACGCACAAACGTCGTA  75
               ||||||||| |||||||||||||||  ||||||||||||||||||||||||||||||||||
Sbjct  4182773  AGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTA  4182832

Query  76       TCTCCGCACTCGGCCCAGGC-TTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  134
               |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Sbjct  4182833  TCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  4182892

Query  135      TACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCG  194
               ||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182893  TACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCG  4182952

Query  195      GTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATAC  238
               |||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  4182953  GTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/10 AF II

Sequence: 50401499

Samples:	16301	Quality: 0 - 9	
Bases:	239	10 - 19	
Average spacing:	69.0	20 - 29	
Average quality >=	10: 114, 20: 73, 30: 28	>= 30	



2019/10 BK

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
315 bits(170)	1e-86	201/216(93%)	2/216(0%)	Plus/Plus

```
Query  17       GCTAGGTGTCTCAGTTTATGAACCAGAAC-CTGCGTTGTCTGAGATTACGCACAAACGTC  75
               || || ||||||||||||||| ||||||||    || ||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  76       GTATCTCCGCACTCGAGCCAGGC-TTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  134
               ||||||||||||||  ||||||  ||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  135      ACGTACACCCTACTCACTACGGCCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAACA  194
               |||||||||| |||||||||| |||||||||| ||||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  195      TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGa  230
               ||||||||||||||||||||||||||||||||||||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

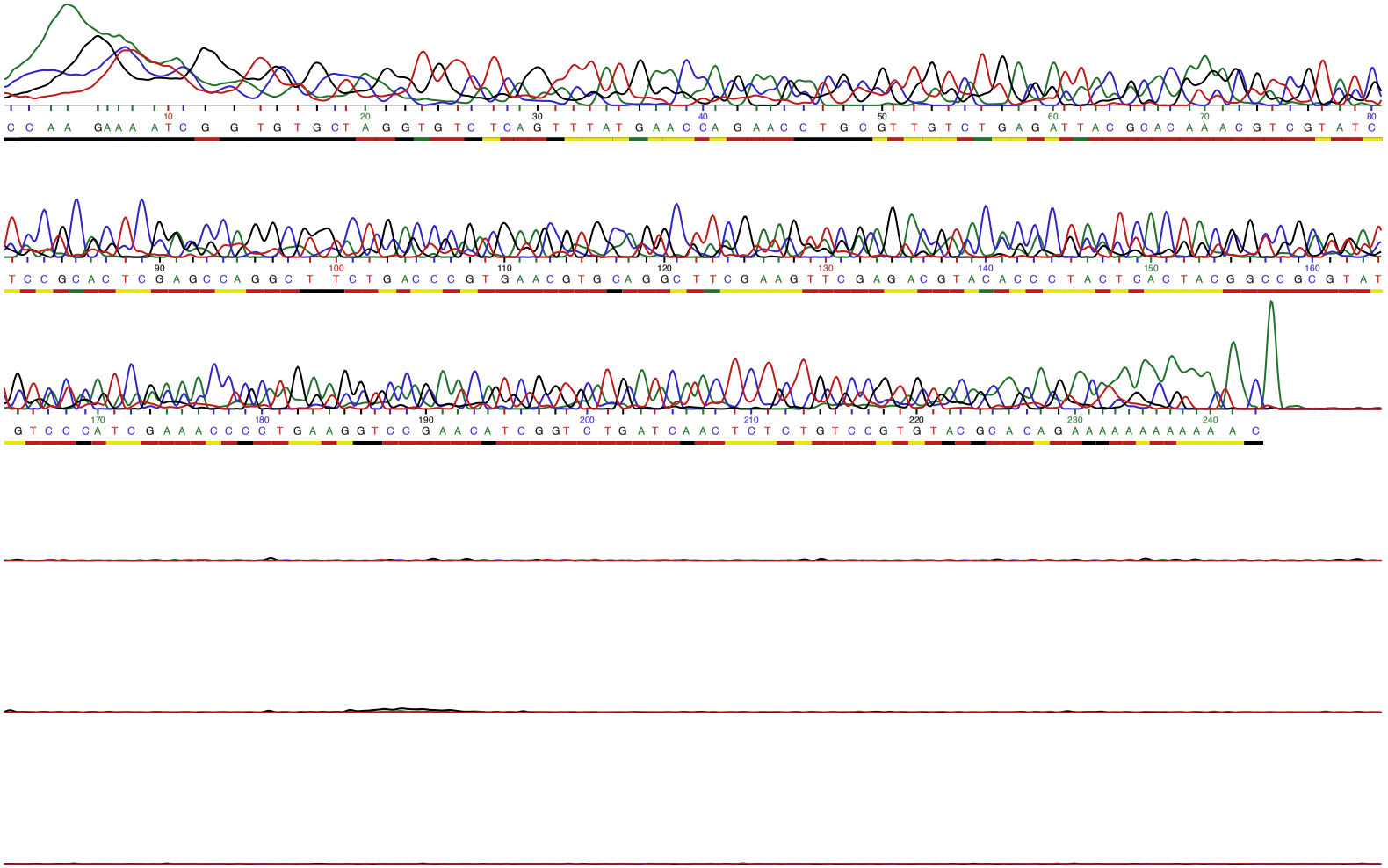
Sequence: 50401543

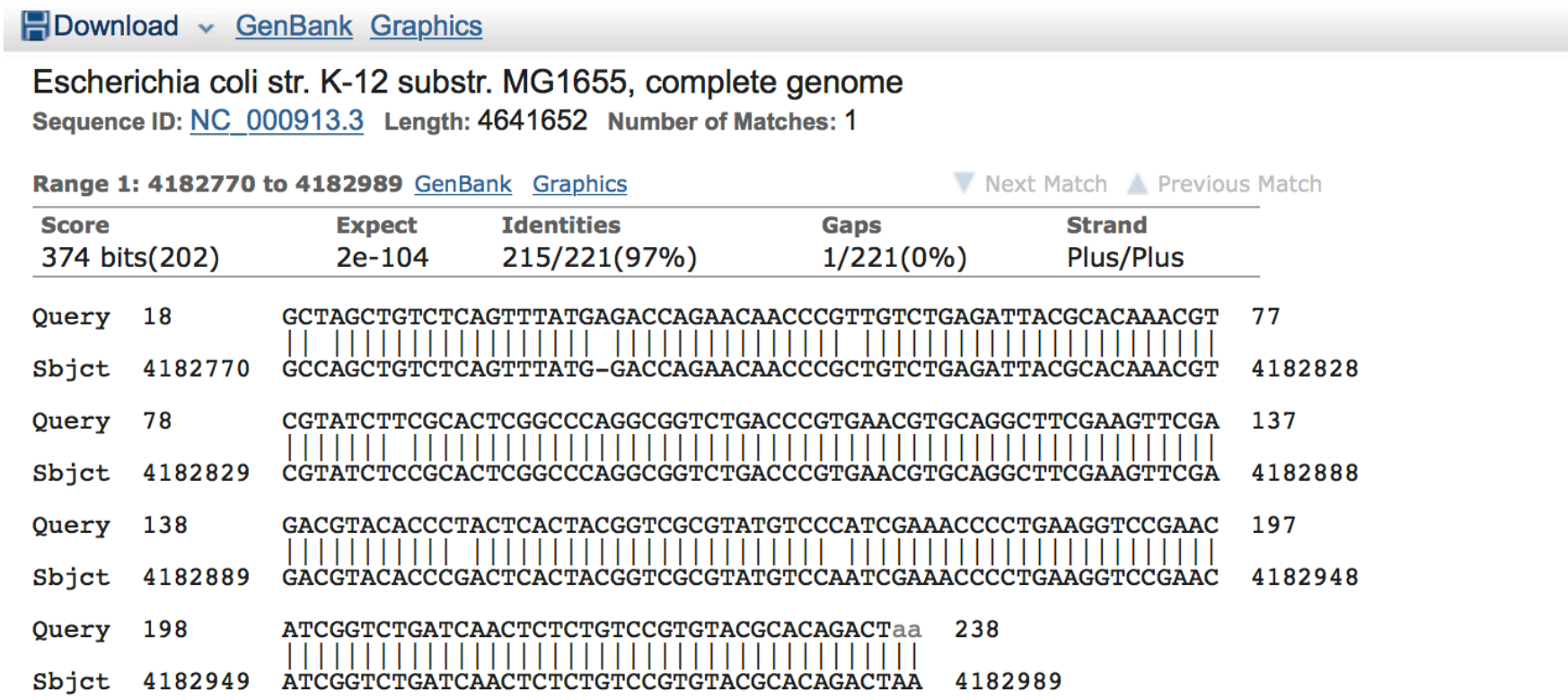
2019/10 BK

Samples: 14813
Bases: 243
Average spacing: 61.0
Average quality >= 10: 129, 20: 67, 30: 7

Quality: 0 - 9
10 - 19
20 - 29
>= 30

29.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182989 GenBank Graphics Next Match Previous Match

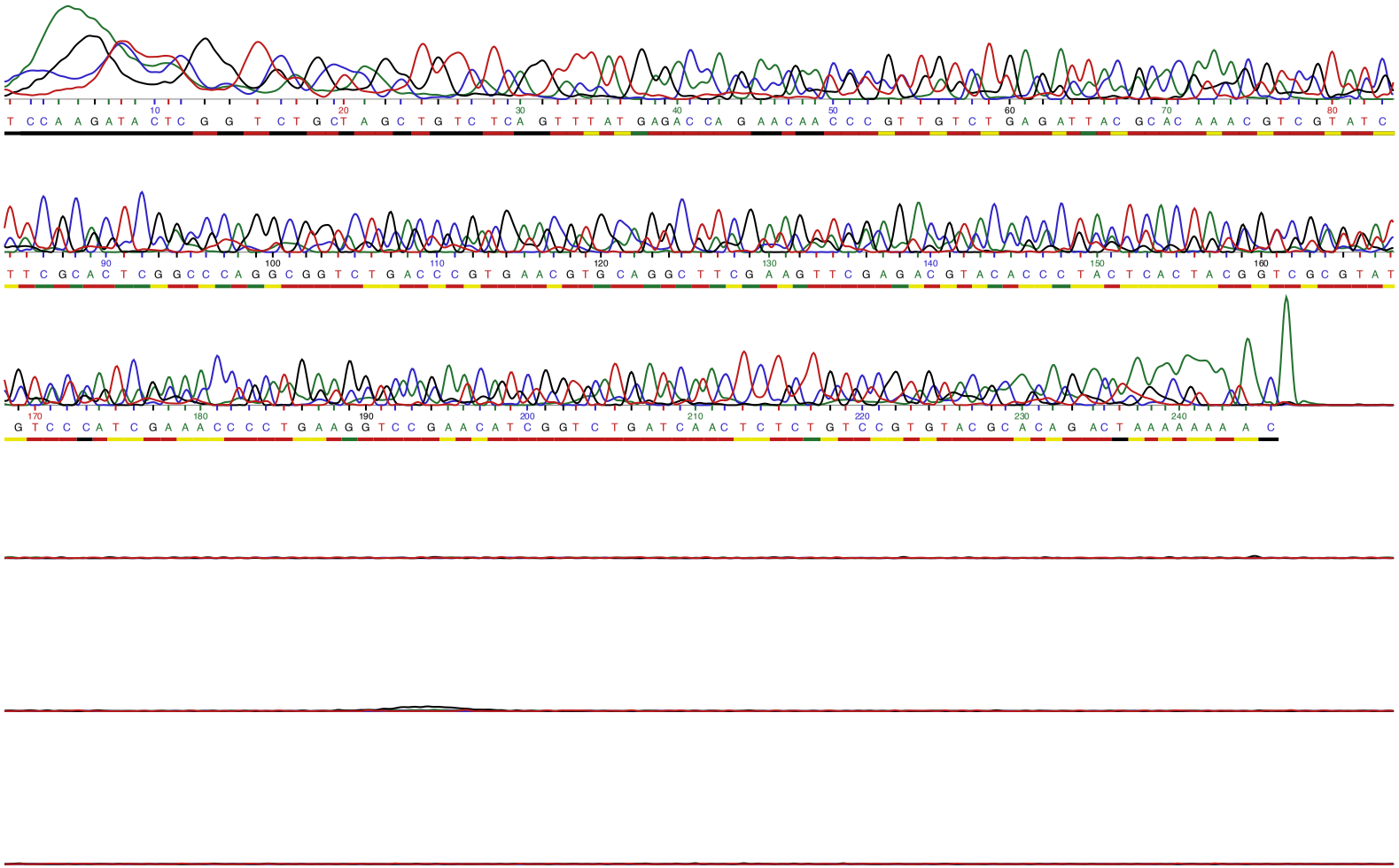
Score	Expect	Identities	Gaps	Strand
374 bits(202)	2e-104	215/221(97%)	1/221(0%)	Plus/Plus

```
Query  18       GCTAGCTGTCTCAGTTTATGAGACCAGAACAACCCGTTGTCTGAGATTACGCACAAACGT  77
               || |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  4182828

Query  78       CGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  137
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182829  CGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  4182888

Query  138      GACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAAC  197
               |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  4182889  GACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAAC  4182948

Query  198      ATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  238
               ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182949  ATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/10 BF I

Sequence: 50401536

Samples:	16302	Quality:	0 - 9
Bases:	246		10 - 19
Average spacing:	67.0		20 - 29
Average quality >=	10: 136, 20: 61, 30: 19		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182773 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
377 bits(204)	3e-105	210/213(99%)	0/213(0%)	Plus/Plus

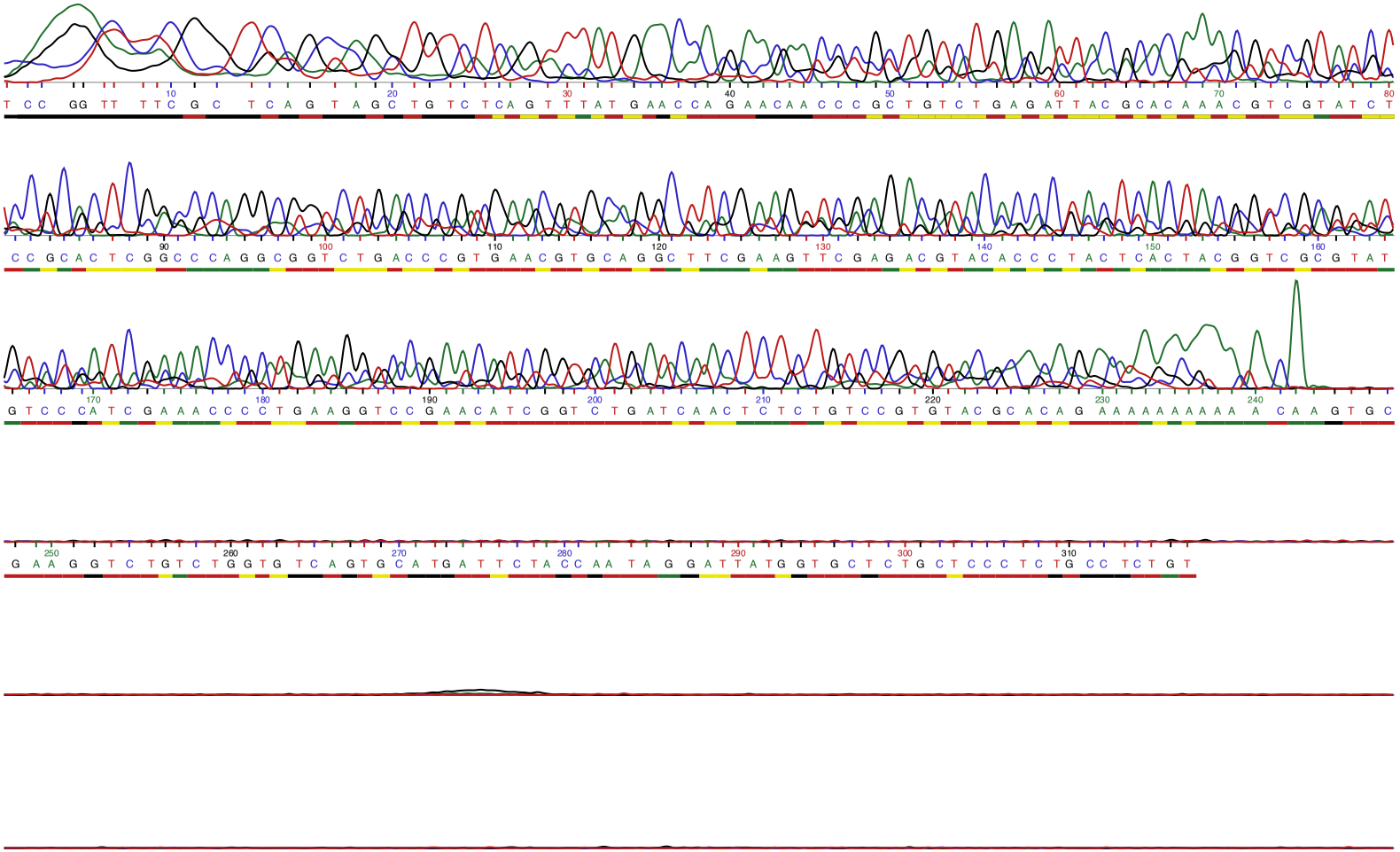
```
Query  18       AGCTGTCTCAGTTTATGAACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTA  77
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182773  AGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTA  4182832

Query  78       TCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  137
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182833  TCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  4182892

Query  138      TACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAACATCG  197
                ||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||
Sbjct  4182893  TACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCG  4182952

Query  198      GTCTGATCAACTCTCTGTCCGTGTACGCACAGa  230
                |||||||||||||||||||||||||||||||
Sbjct  4182953  GTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2019/10 BF II

Sequence: 50401529

Samples:	16300	Quality:	0 - 9
Bases:	318		10 - 19
Average spacing:	52.0		20 - 29
Average quality >=	10: 146, 20: 82, 30: 48		>= 30

29.03.2019



2019/10 CK

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182988 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
377 bits(204)	3e-105	215/220(98%)	1/220(0%)	Plus/Plus

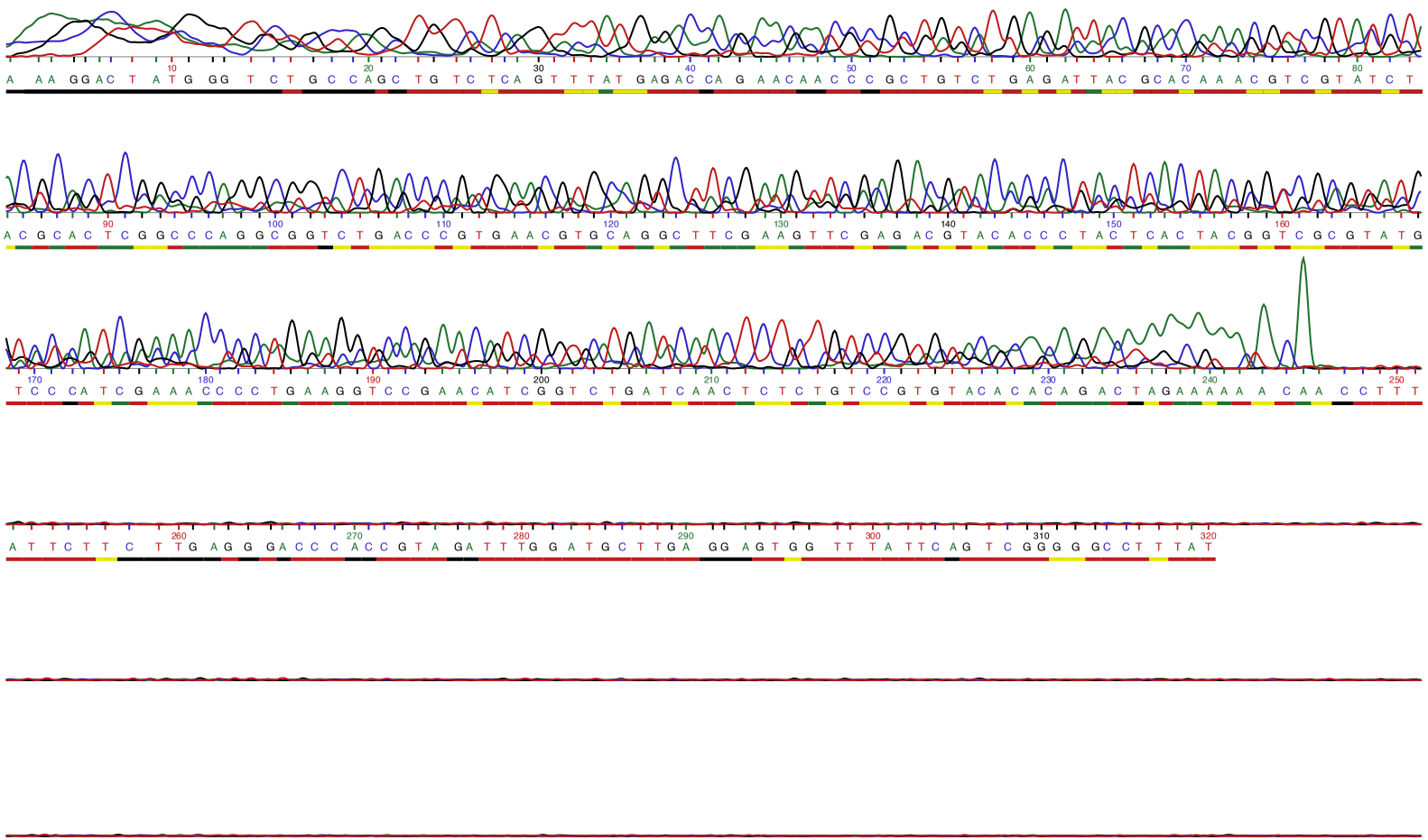
```
Query  17       GCCAGCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  76
               ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  4182828

Query  77       CGTATCTACGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  136
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182829  CGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  4182888

Query  137      GACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAAC  196
               |||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  4182889  GACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAAC  4182948

Query  197      ATCGGTCTGATCAACTCTCTGTCCGTGTACACACAGACTA  236
               |||||||||||||||||||||||||||||| |||||||||
Sbjct  4182949  ATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTA  4182988
```

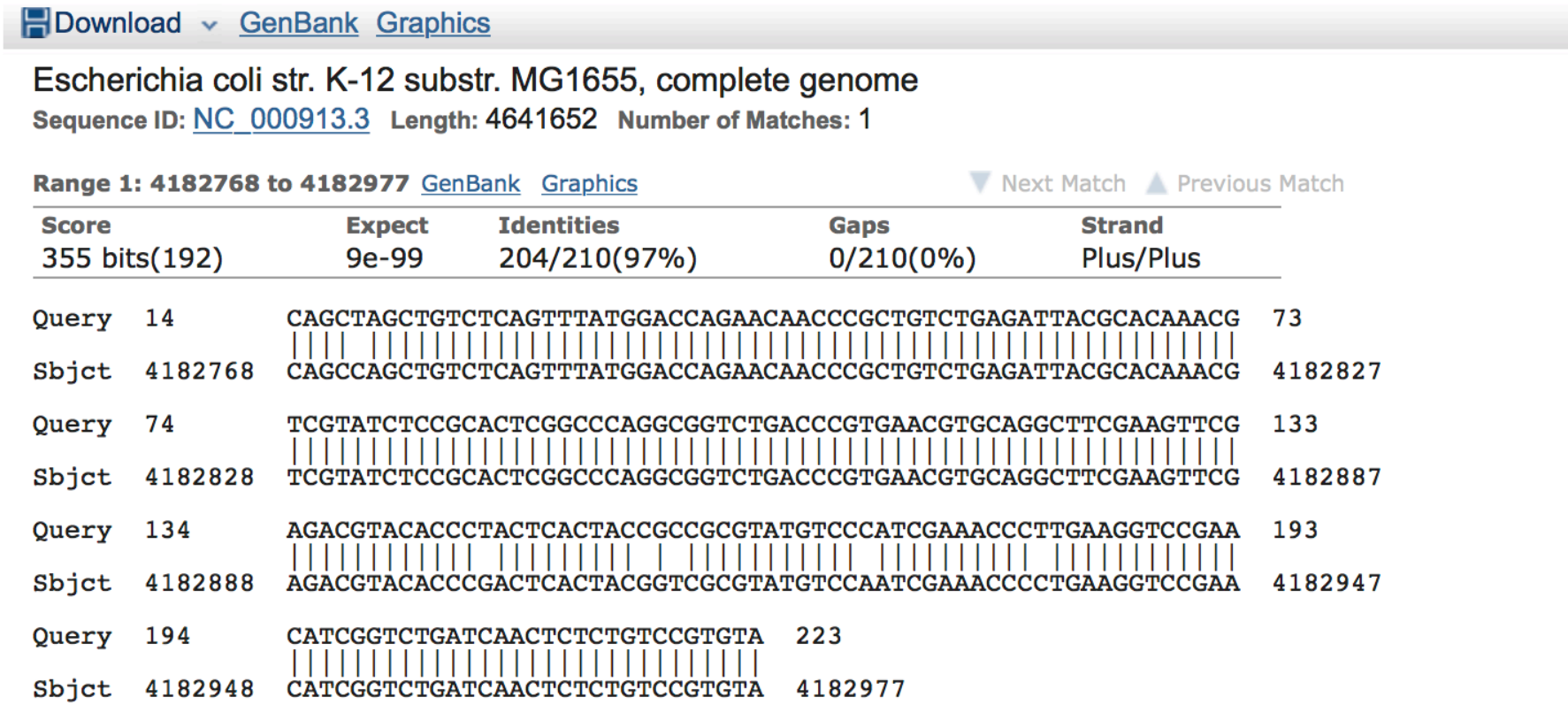
2019/10 CK

Sequence: 50401574

Samples:	16301	Quality:	0 - 9
Bases:	321		10 - 19
Average spacing:	51.0		20 - 29
Average quality >=	10: 170, 20: 69, 30: 39		>= 30

01.04.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182977 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
355 bits(192)	9e-99	204/210(97%)	0/210(0%)	Plus/Plus

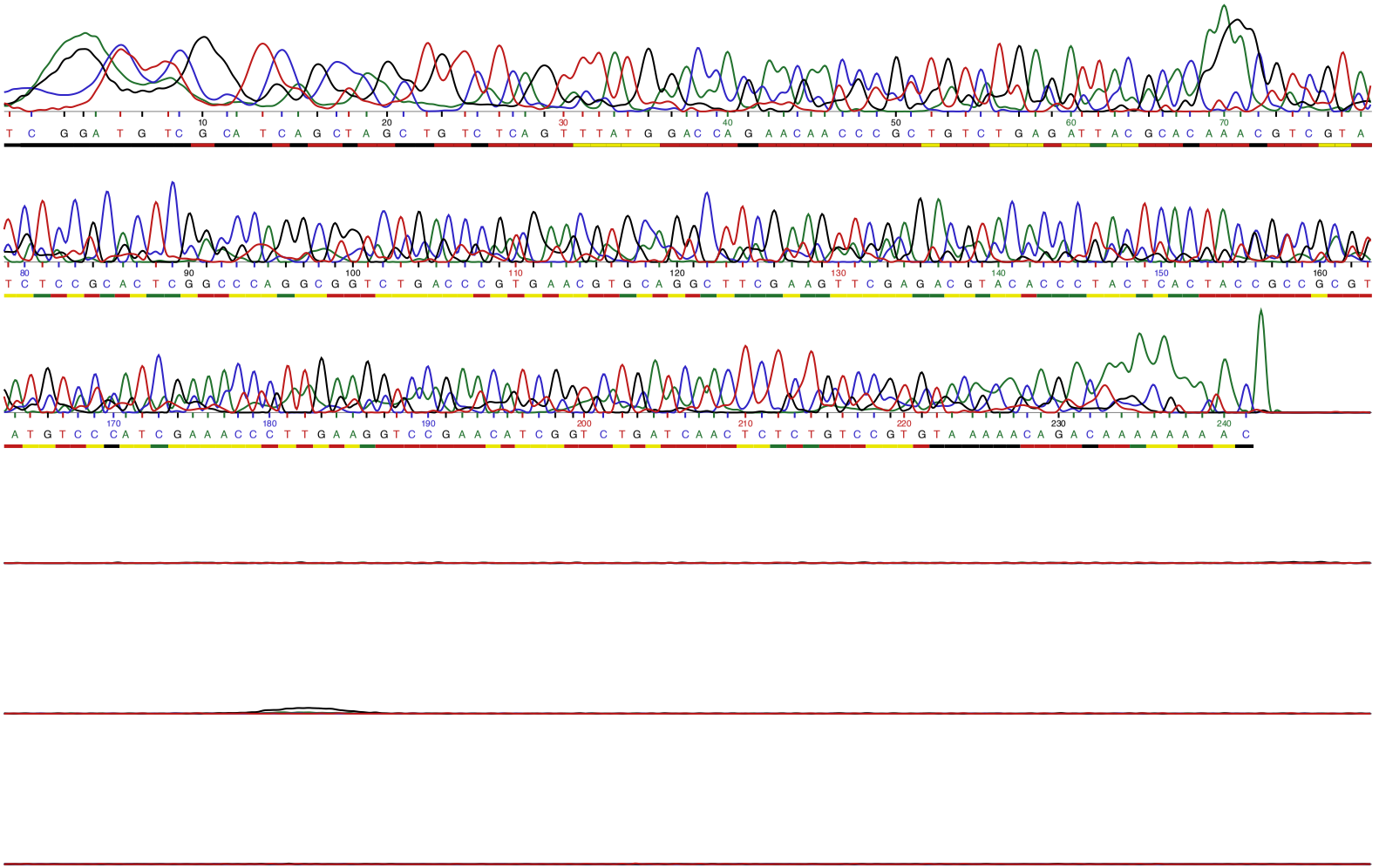
```
Query  14       CAGCTAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  73
                ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  74       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCTACTCACTACCGCCGCGTATGTCCCATCGAAACCCTTGAAGGTCCGAA  193
                |||||||||||| |||||||||| | |||||||||||| |||||||||| ||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGGTCTGATCAACTCTCTGTCCGTGTA  223
                ||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

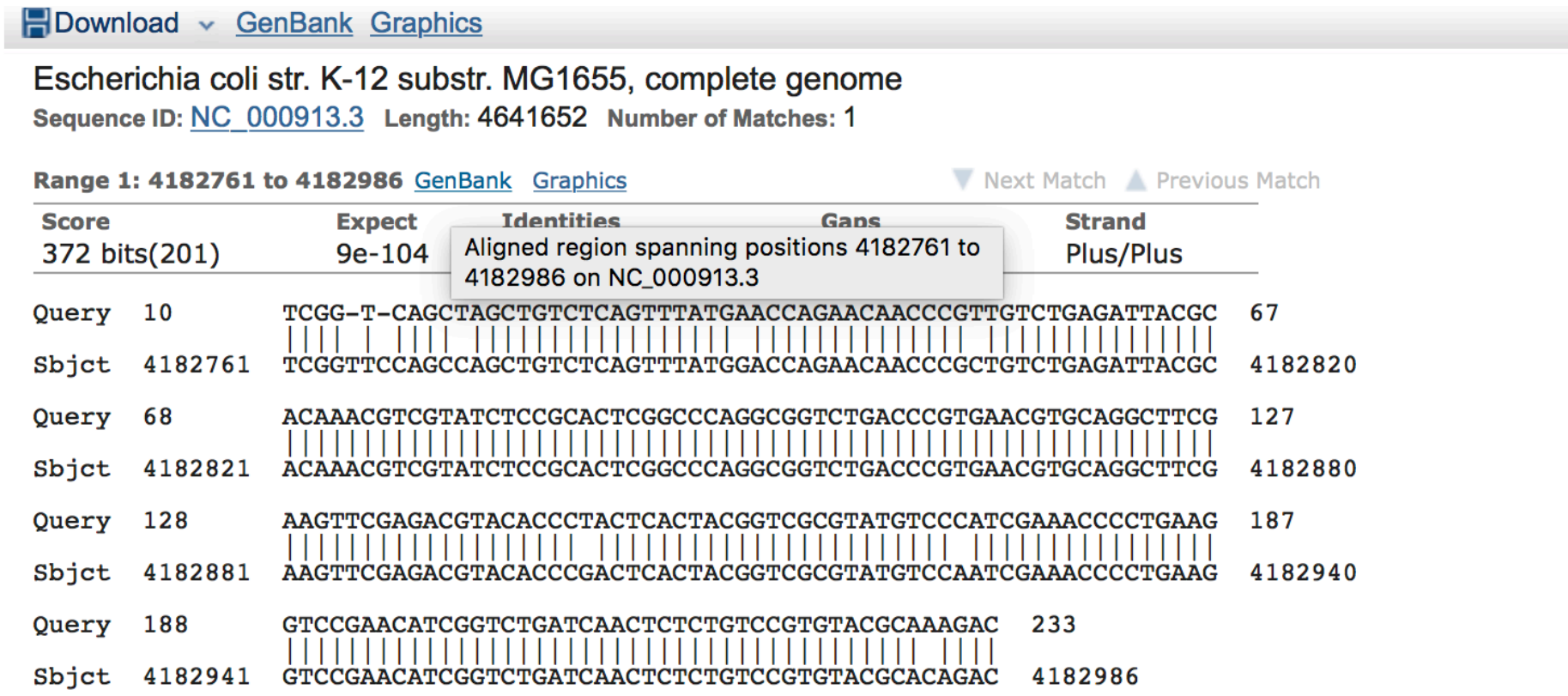
2019/10 DF I

Sequence: 50401604

Samples:	11853	Quality: 0 - 9	
Bases:	242	10 - 19	
Average spacing:	49.0	20 - 29	
Average quality >=	10: 105, 20: 81, 30: 27	>= 30	

Page: 1 / 2
29.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
372 bits(201)	9e-104			Plus/Plus

Aligned region spanning positions 4182761 to 4182986 on NC_000913.3

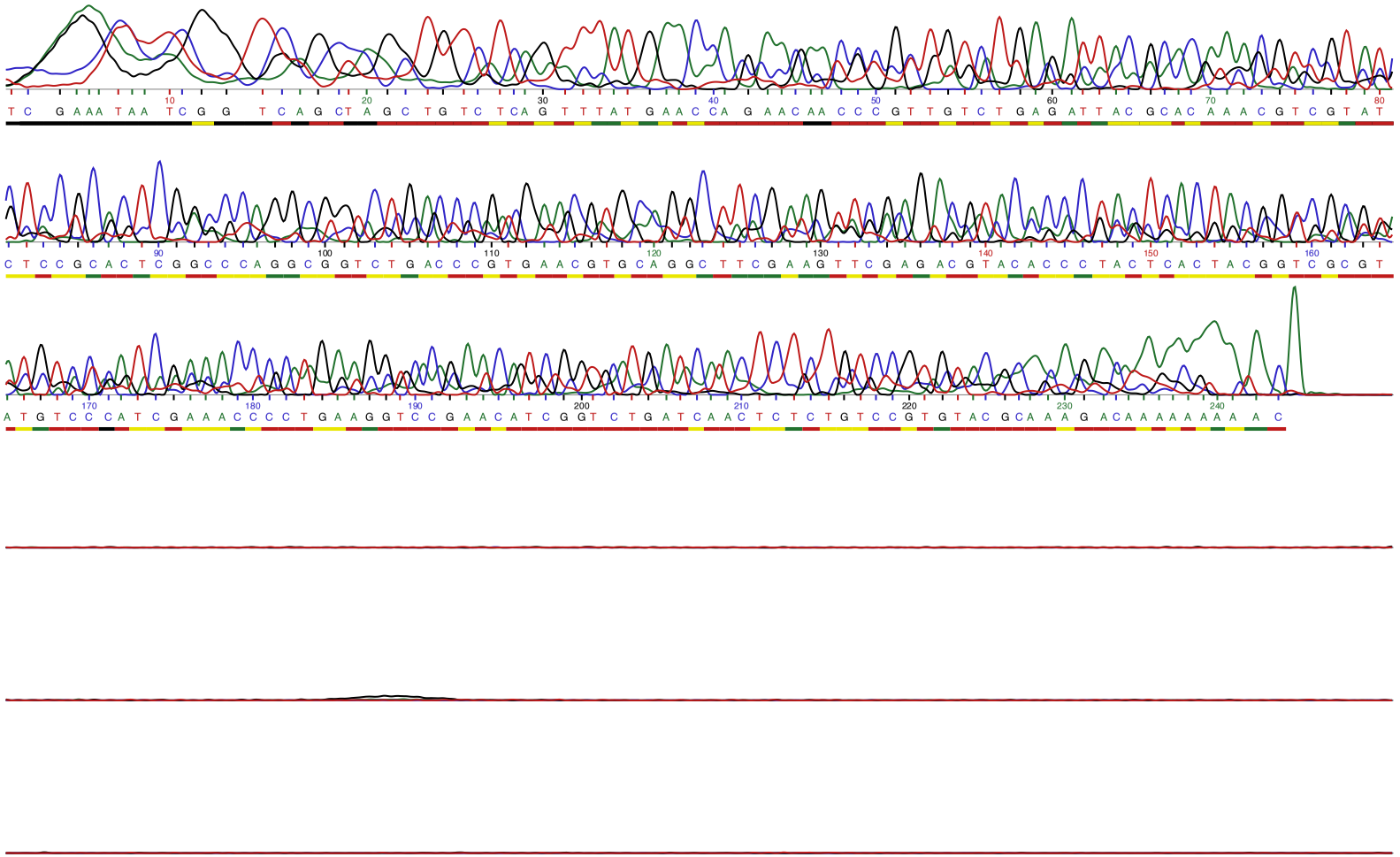
```
Query  10       TCGG-T-CAGCTAGCTGTCTCAGTTTATGAACCAGAACAACCCGTTGTCTGAGATTACGC  67
                ||||| | ||||| |||||||||||||||||| |||||||||||||| ||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  68       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  127
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  128      AAGTTCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAG  187
                |||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  188      GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCAAAGAC  233
                |||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

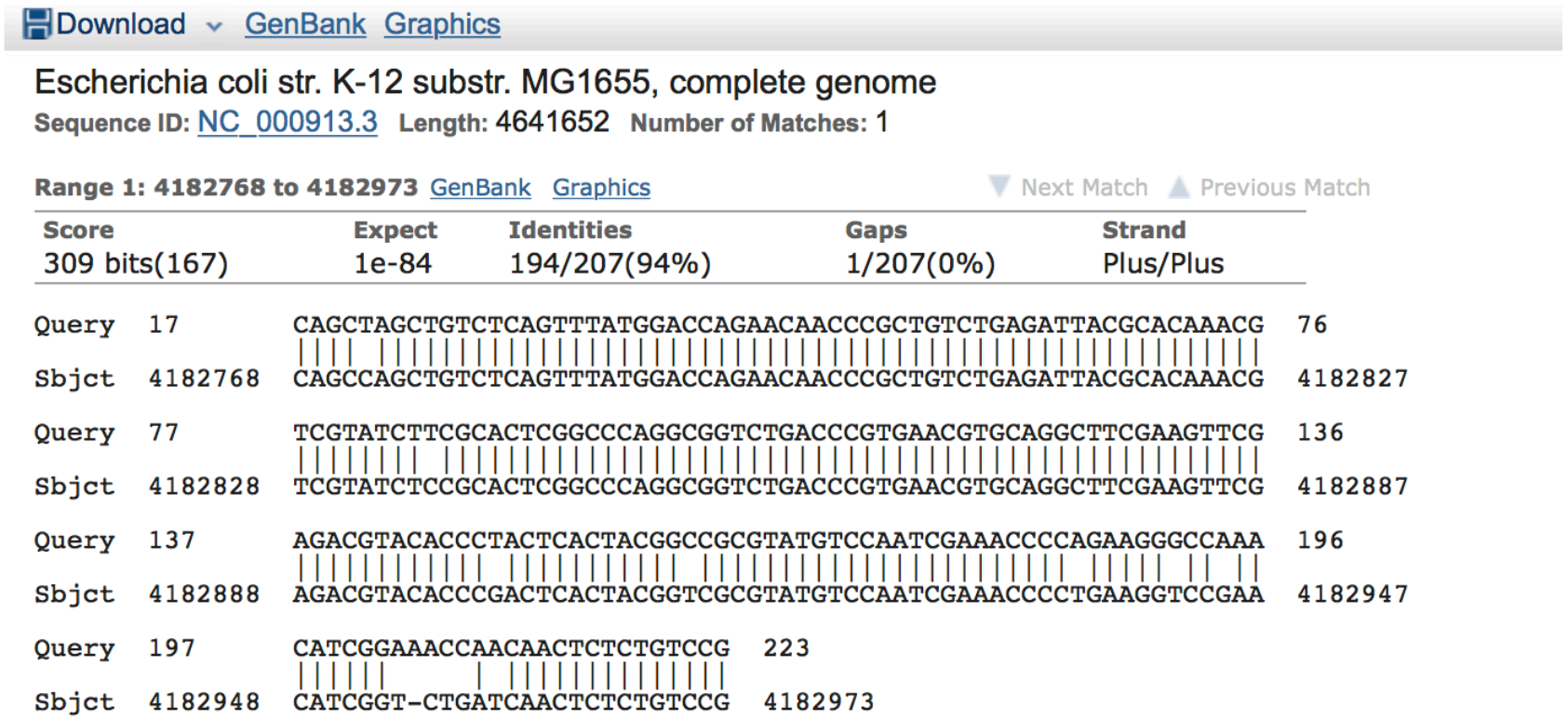
2019/10 DF II

Sequence: 50401598

Samples:	13572	Quality:	0 - 9
Bases:	244		10 - 19
Average spacing:	56.0		20 - 29
Average quality >=	10: 117, 20: 81, 30: 27		>= 30

29.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182973 GenBank Graphics Next Match Previous Match

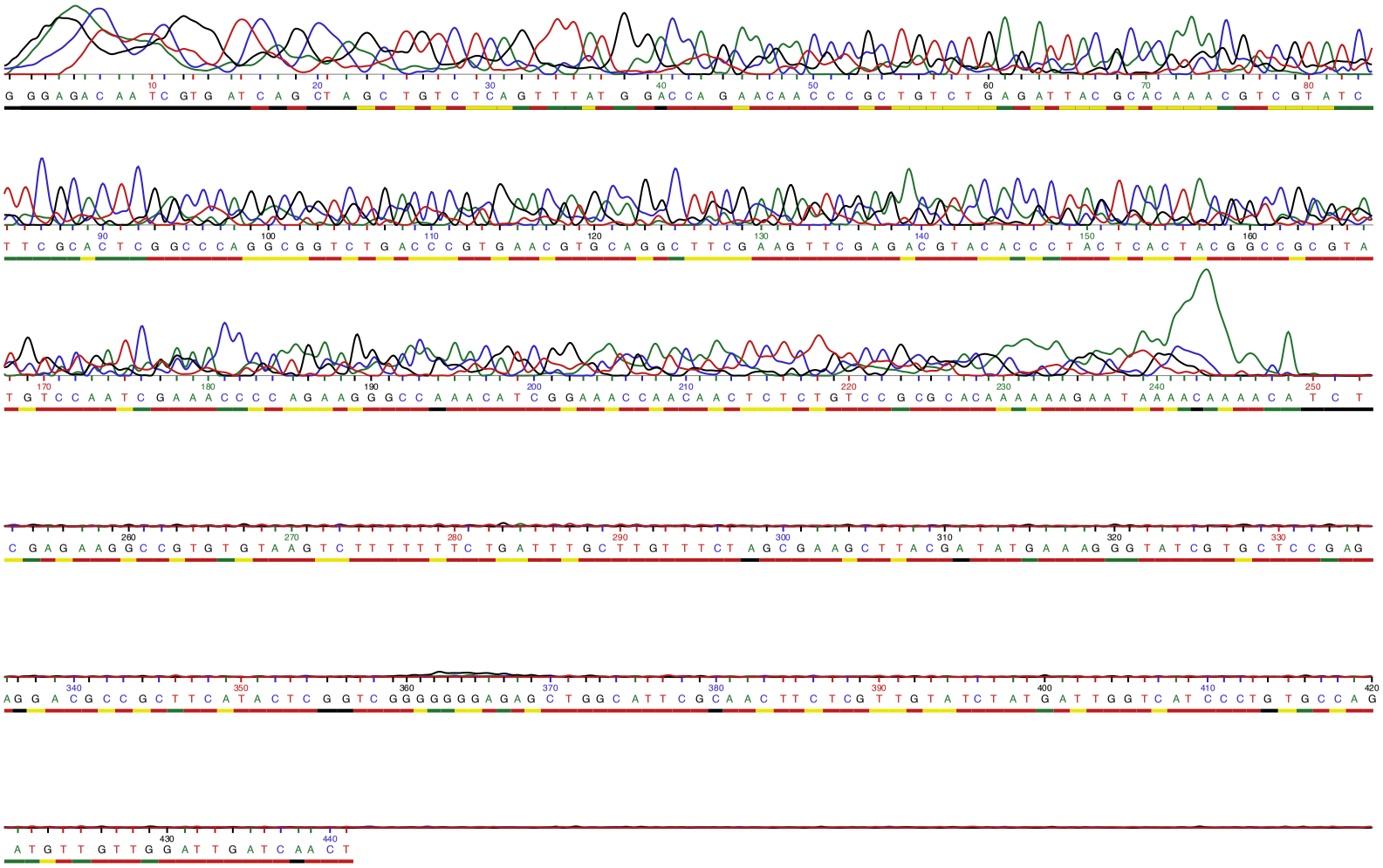
Score	Expect	Identities	Gaps	Strand
309 bits(167)	1e-84	194/207(94%)	1/207(0%)	Plus/Plus

```
Query  17       CAGCTAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  76
                ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  77       TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  136
                ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  137      AGACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCAGAAGGGCCAAA  196
                ||||||||||||  ||||||||||  ||||||||||||||||||||||  |||||  ||  ||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  197      CATCGGAAACCAACAACTCTCTGTCCG  223
                ||||||      | ||||||||||||||
Sbjct  4182948  CATCGGT-CTGATCAACTCTCTGTCCG  4182973
```

2019/11 AF II

Sequence: 50401635

Samples:	16303	Quality:	0 - 9
Bases:	442		10 - 19
Average spacing:	37.0		20 - 29
Average quality >=	10: 257, 20: 107, 30: 44		>= 30



2019/11 BK

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182975 GenBank Graphics Next Match Previous Match

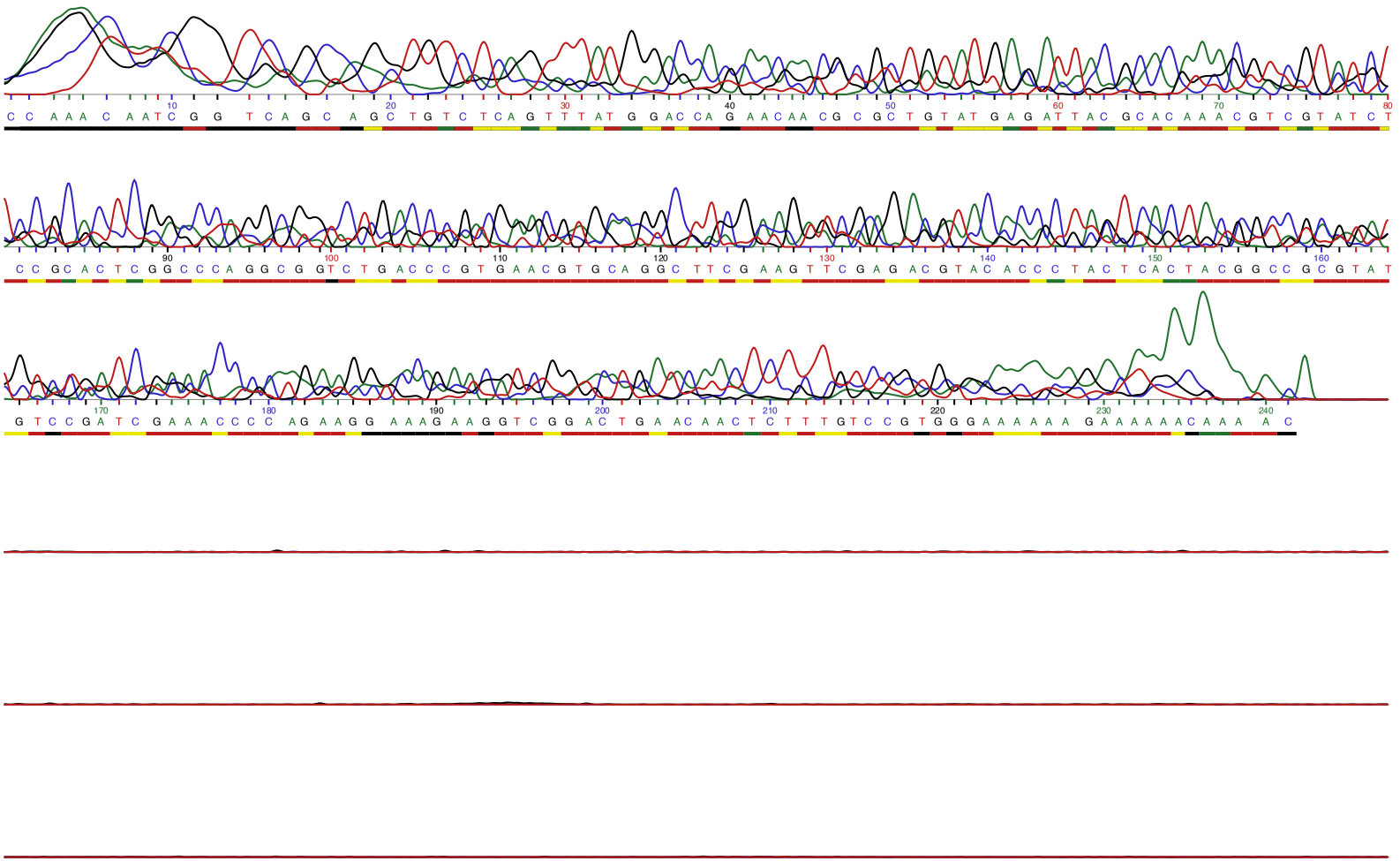
Score	Expect	Identities	Gaps	Strand
300 bits(162)	4e-82	198/215(92%)	3/215(1%)	Plus/Plus

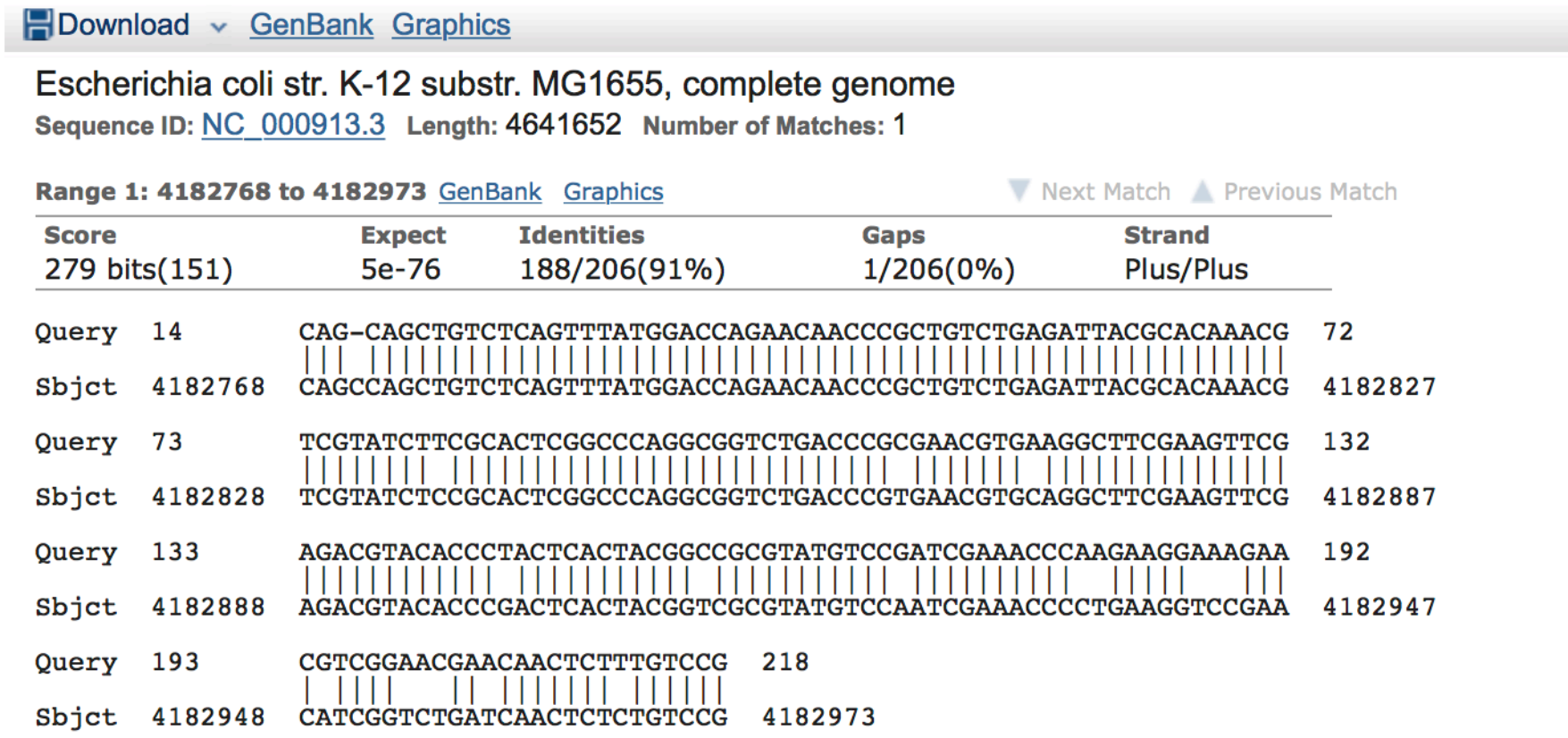
```
Query  9        TCGG-T-CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACGCGCTGTATGAGATTACGC  65
               ||||  |  |||  ||||||||||||||||||||||||||||||||  ||||||  ||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGC  4182820

Query  66       ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182821  ACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCG  4182880

Query  126      AAGTTCGAGACGTACACCCTACTCACTACGGCCGCGTATGTCCGATCGAAACCCCAGAAG  185
               ||||||||||||||||||  ||||||||||  ||||||||||  ||||||||||  ||||
Sbjct  4182881  AAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAG  4182940

Query  186      GAAAGAAGGTCGGACTGAACAACTCTTTGTCCGTG  220
               |    |||   ||||  ||||  ||||||||  |||||||||
Sbjct  4182941  GTCCGAACATCGGTCTGATCAACTCTCTGTCCGTG  4182975
```

2019/11 BK

Sequence: 50401628

Samples:	16303	Quality:	0 - 9
Bases:	242		10 - 19
Average spacing:	68.0		20 - 29
Average quality >=	10: 134, 20: 60, 30: 16		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182973 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
279 bits(151)	5e-76	188/206(91%)	1/206(0%)	Plus/Plus

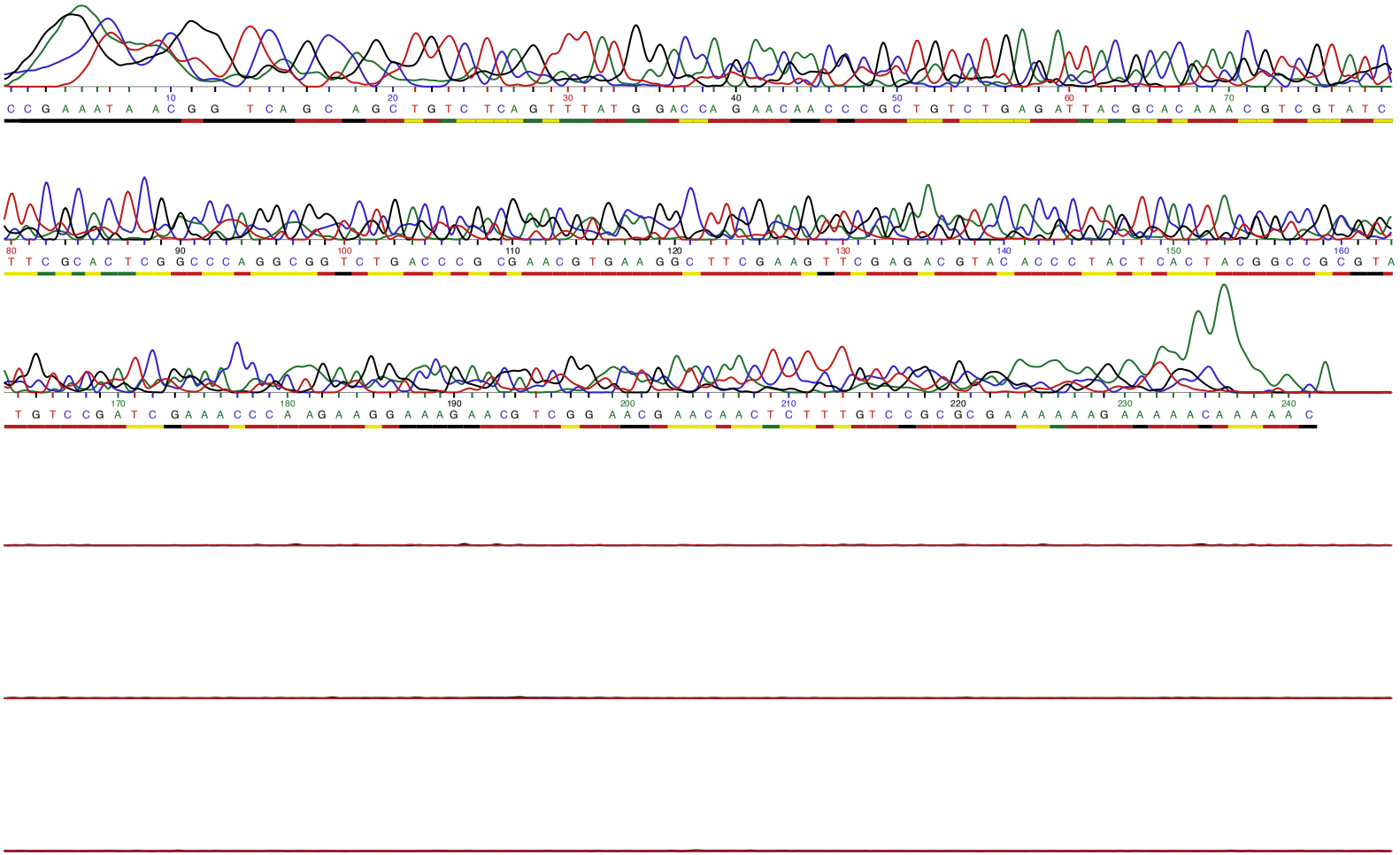
```
Query  14       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  72
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  73       TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGCGAACGTGAAGGCTTCGAAGTTCG  132
               |||||||| ||||||||||||||||||||||||||| ||||||| ||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  133      AGACGTACACCCTACTCACTACGGCCGCGTATGTCCGATCGAAACCCAAGAAGGAAAGAA  192
               |||||||||||| |||||||||||| |||||||||||| ||||||||||   |||||   |||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  193      CGTCGGAACGAACAACTCTTTGTCCG  218
               | ||||   || ||||||| ||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCG  4182973
```

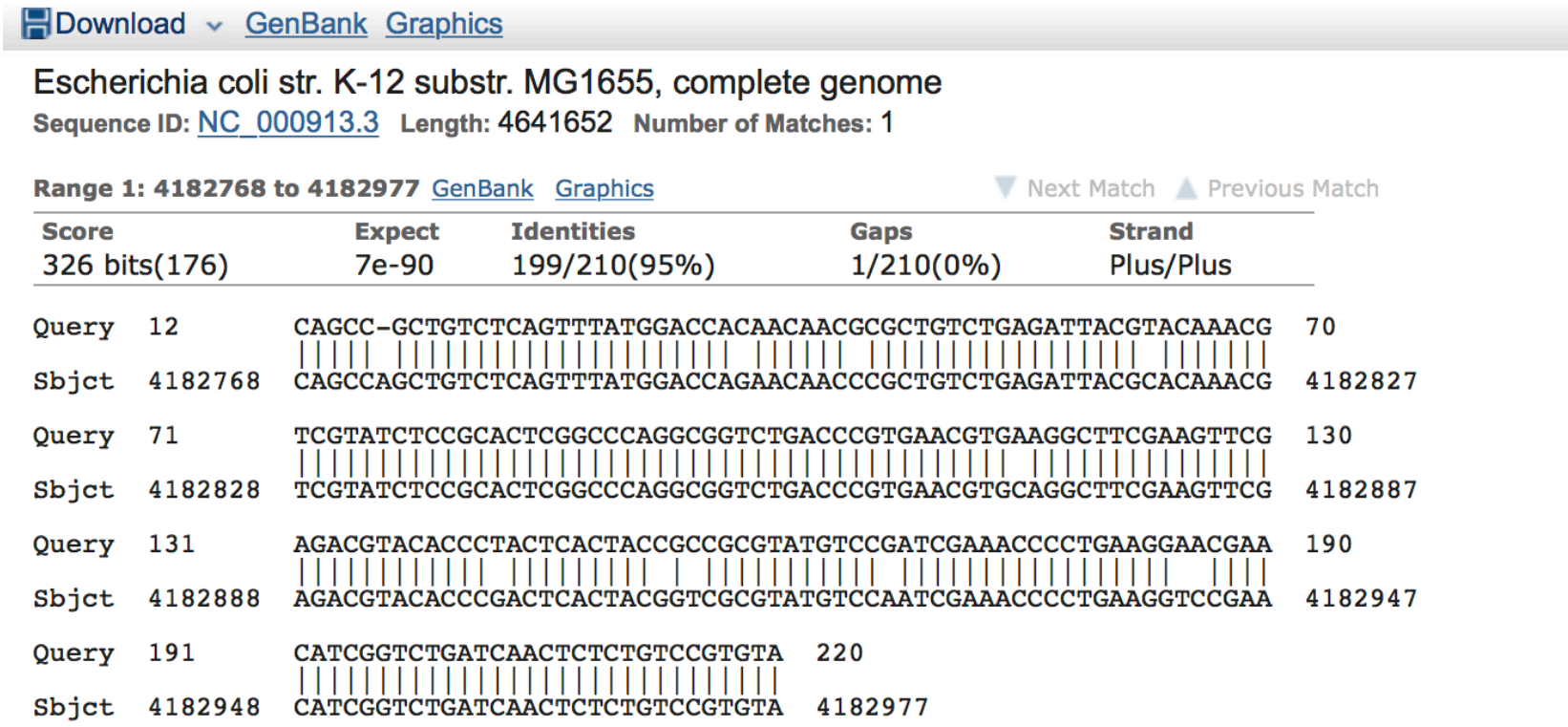
2019/11 BF I

Sequence: 50401611

Samples:	16301	Quality:	0 - 9
Bases:	242		10 - 19
Average spacing:	68.0		20 - 29
Average quality >=	10: 125, 20: 69, 30: 13		>= 30

29.03.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182977 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
326 bits(176)	7e-90	199/210(95%)	1/210(0%)	Plus/Plus

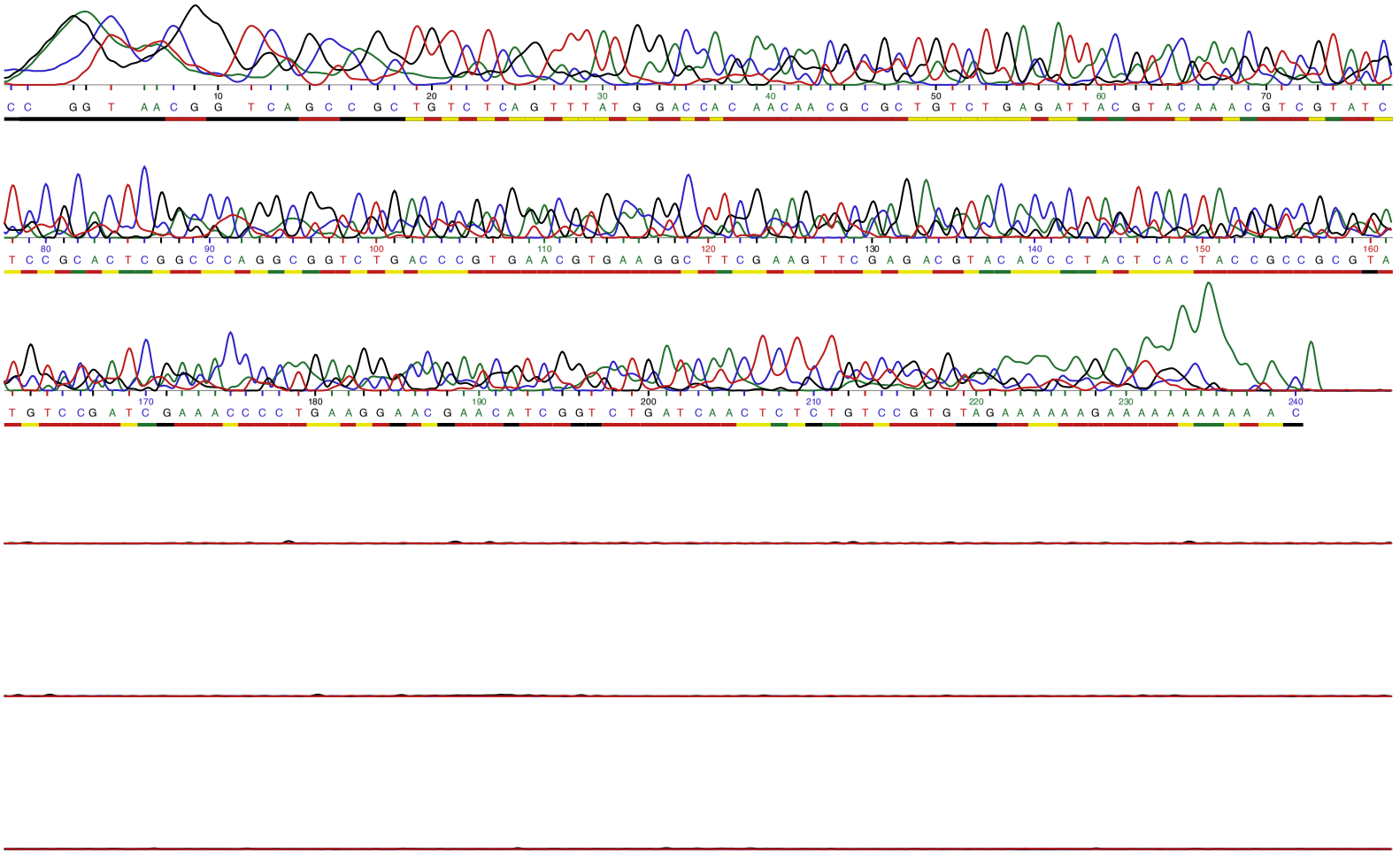
```
Query  12       CAGCC-GCTGTCTCAGTTTATGGACCACAACAACGCGCTGTCTGAGATTACGTACAAACG  70
                |||||  ||||||||||||||||||||| |||||| |||||||||||||||||| |||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  71       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGAAGGCTTCGAAGTTCG  130
                ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  131      AGACGTACACCCTACTCACTACCGCCGCGTATGTCCGATCGAAACCCCTGAAGGAACGAA  190
                |||||||||||| ||||||||| | |||||||||| |||||||||||||||||   ||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  191      CATCGGTCTGATCAACTCTCTGTCCGTGTA  220
                |||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

2019/11 DF II

Sequence: 50401666

Samples:	16300	Quality:	0 - 9
Bases:	241		10 - 19
Average spacing:	68.0		20 - 29
Average quality >=	10: 125, 20: 70, 30: 19		>= 30

29.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182996 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
372 bits(201)	2e-103	220/229(96%)	2/229(0%)	Plus/Plus

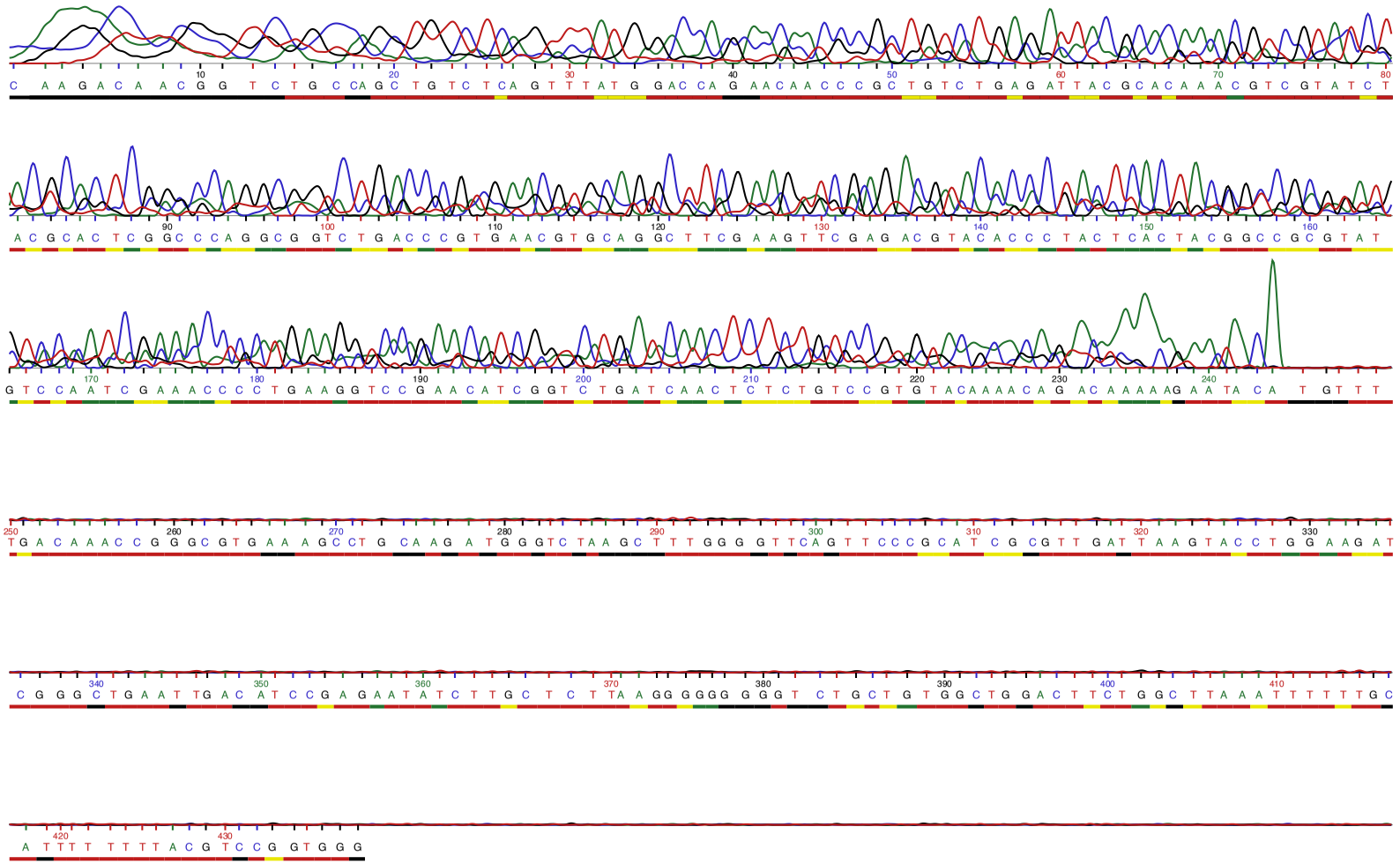
```
Query  15       GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  74
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  75       GTATCTACGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  134
                ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  135      ACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  194
                ||||||||||  ||||||||||  ||||||||||||||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  195      TCGGTCTGATCAACTCTCTGTCCGTGTACAAAACAGACAAAAAGAATAC  243
                |||||||||||||||||||||||||||||  | |||||  ||  ||||||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTACGCA-CAGACTAAC-GAATAC  4182996
```

Sequence: 29415908

2019/15 AK I*

Samples:	16301	Quality:	0 - 9
Bases:	439		10 - 19
Average spacing:	38.0		20 - 29
Average quality >=	10: 250, 20: 85, 30: 50		>= 30

22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182953 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
237 bits(128)	6e-63	168/187(90%)	3/187(1%)	Plus/Plus

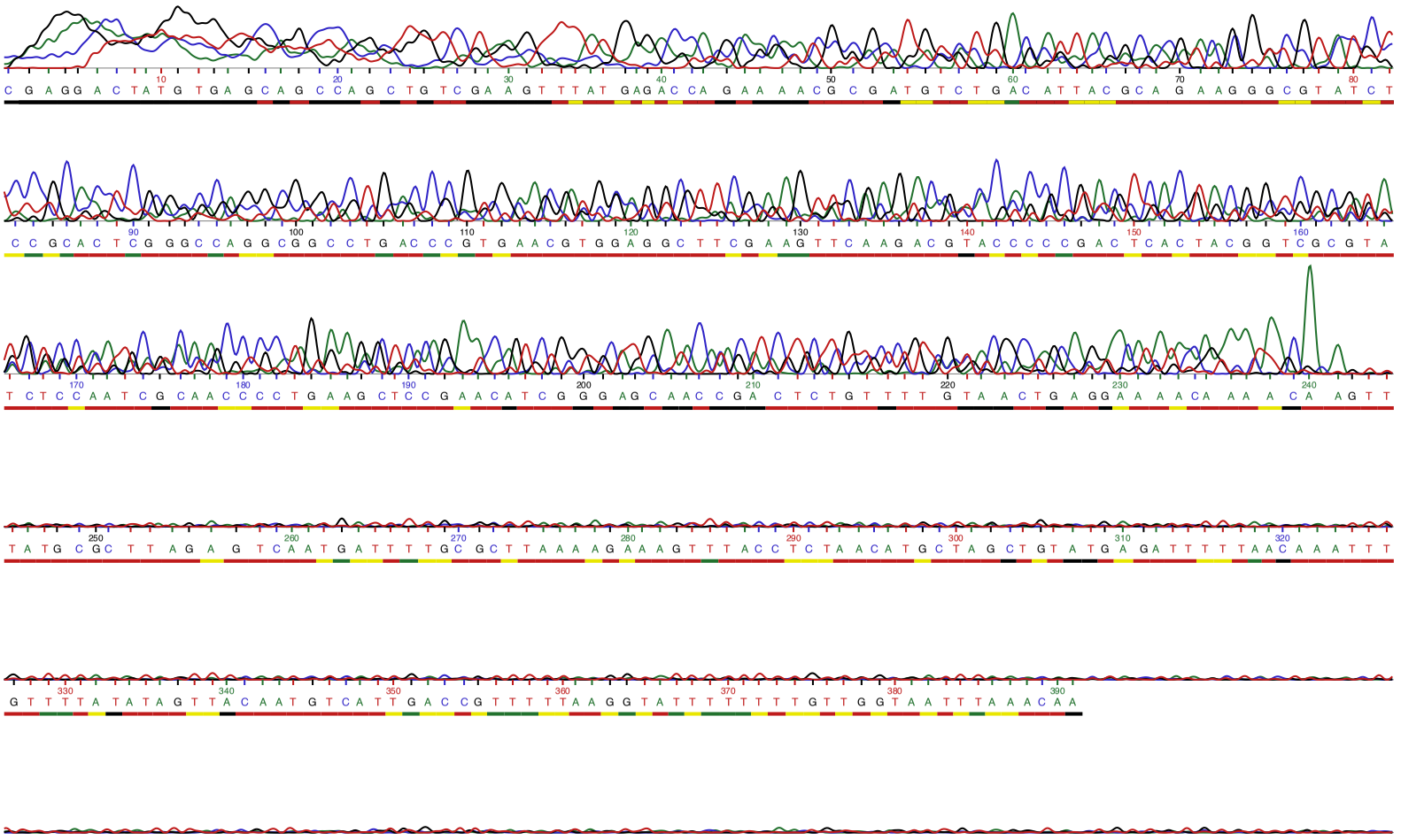
```
Query  16       CAGCCAGCTGTCGAAGTTTATGAGACCAGAA-AACGCGATGTCTGACATTACGCA-GAAG  73
                ||||||||||||  |||||||| |||||||| ||| || ||||||| ||||||||  ||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  74       GGCGTATCTCCGCACTCGGGCCAGGCGGCCTGACCCGTGAACGTGGAGGCTTCGAAGTTC  133
                | |||||||||||||||| |||||||| |||||||||||||| |||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  134      AAGACGTACCCCCGACTCACTACGGTCGCGTATCTCCAATCGCAACCCCTGAAGCTCCGA  193
                 |||||||| |||||||||||||||||||||||| |||||||| |||||||||| |||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  194      ACATCGG  200
                |||||||
Sbjct  4182947  ACATCGG  4182953
```

2019/15 AK III*

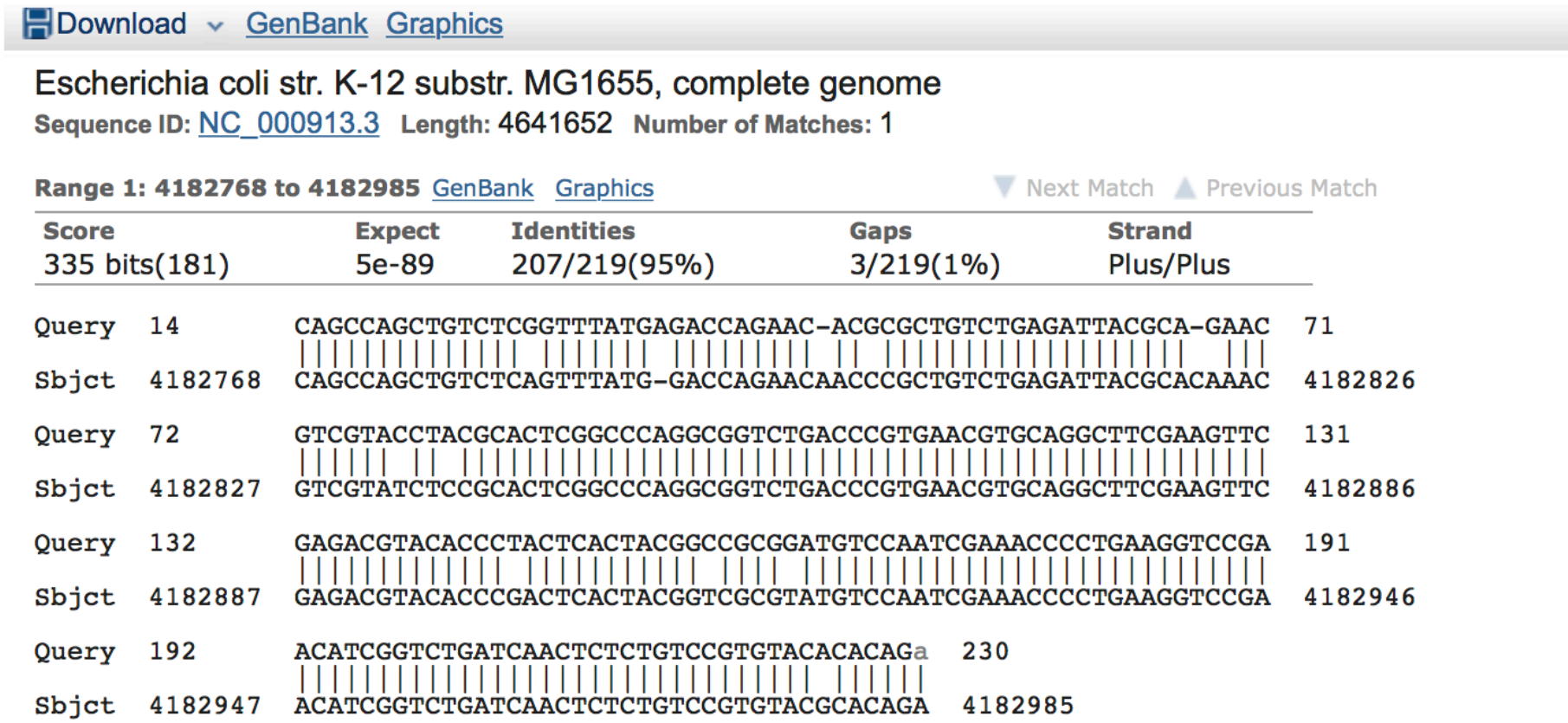
Sequence: 294153

Samples:	16300	Quality: 0 - 9	
Bases:	392	10 - 19	
Average spacing:	42.0	20 - 29	
Average quality >=	10: 233, 20: 76, 30: 27	>= 30	

22.05.2019



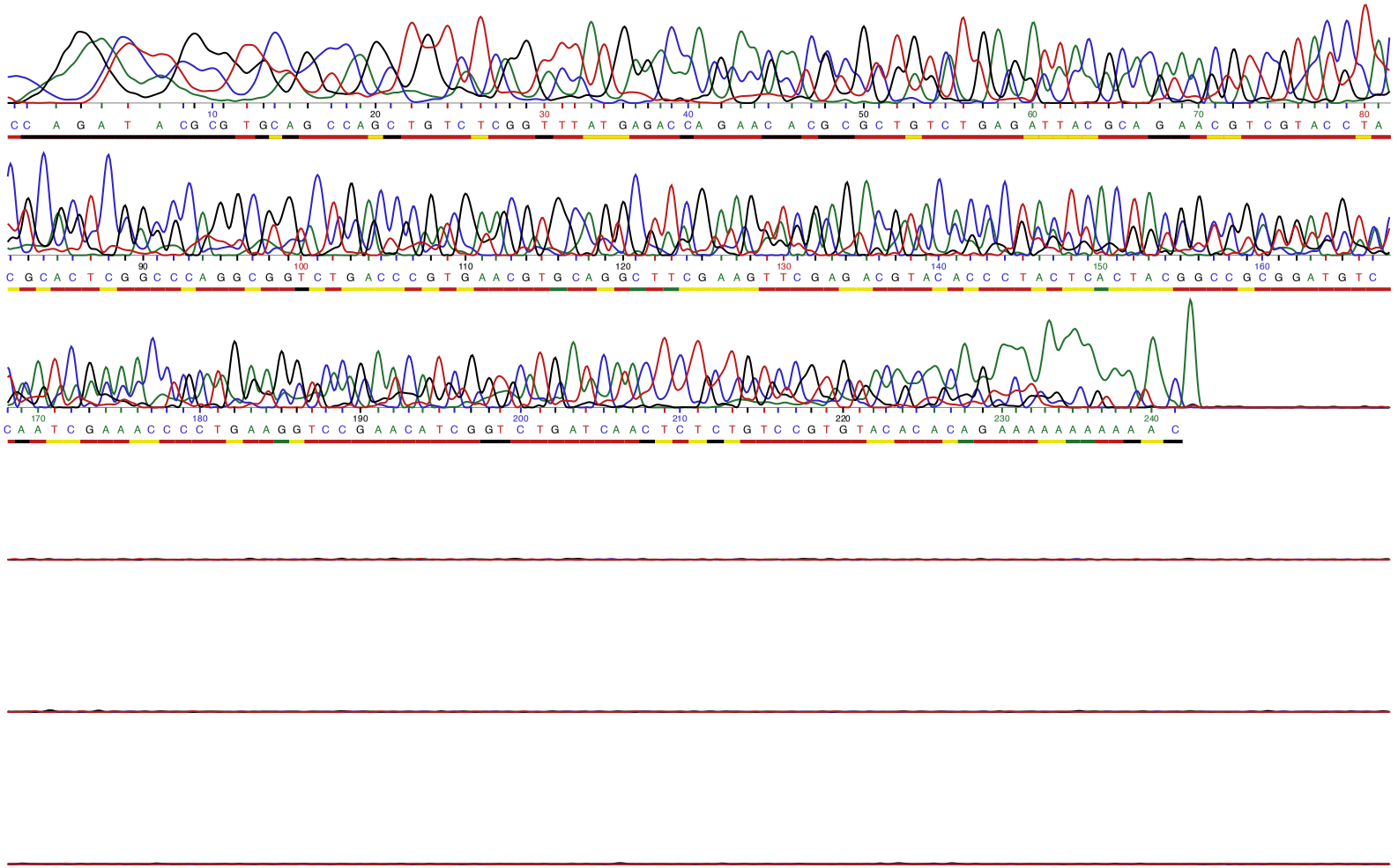
2019/15 BK I*


Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182768 to 4182985 GenBank Graphics
Next Match Previous Match
Score 335 bits(181)
Expect 5e-89
Identities 207/219(95%)
Gaps 3/219(1%)
Strand Plus/Plus
Query 14 CAGCCAGCTGTCTCGGTTTATGAGACCAGAAC-ACGCGCTGTCTGAGATTACGCA-GAAC 71
Sbjct 4182768 CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC 4182826
Query 72 GTCGTACCTACGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC 131
Sbjct 4182827 GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC 4182886
Query 132 GAGACGTACACCCTACTCACTACGGCCGCGGATGTCCAATCGAAACCCCTGAAGGTCCGA 191
Sbjct 4182887 GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA 4182946
Query 192 ACATCGGTCTGATCAACTCTCTGTCCGTGTACACACAGa 230
Sbjct 4182947 ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA 4182985

2019/15 BK I*

Sequence: 29415946

Samples:	15541	Quality: 0 - 9	
Bases:	242	10 - 19	
Average spacing:	65.0	20 - 29	
Average quality >=	10: 145, 20: 60, 30: 8	>= 30	



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182773 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
329 bits(178)	3e-87	202/213(95%)	4/213(1%)	Plus/Plus

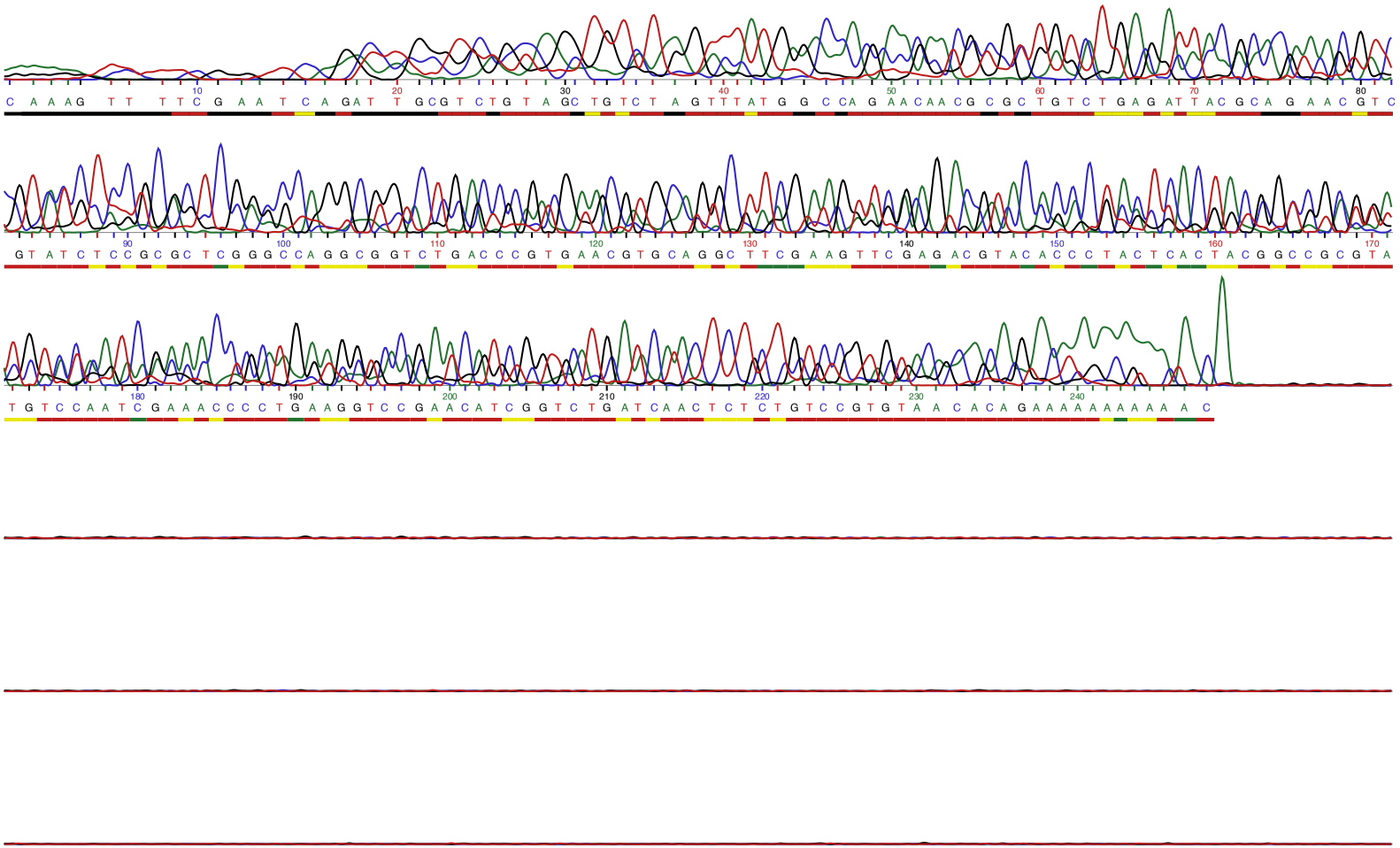
```
Query  29       AGCTGTCT-AGTTTATGG-CCAGAACAACGCGCTGTCTGAGATTACGCA-GAACGTCGTA  85
               ||||||||| |||||||||| |||||||||| |||||||||||||||||||  ||||||||||
Sbjct  4182773  AGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTA  4182832

Query  86       TCTCCGCGCTCGGGCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  145
               ||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182833  TCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACG  4182892

Query  146      TACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCG  205
               ||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  4182893  TACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCG  4182952

Query  206      GTCTGATCAACTCTCTGTCCGTGTAA-CACAGa  237
               |||||||||||||||||||||||||  ||||||
Sbjct  4182953  GTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

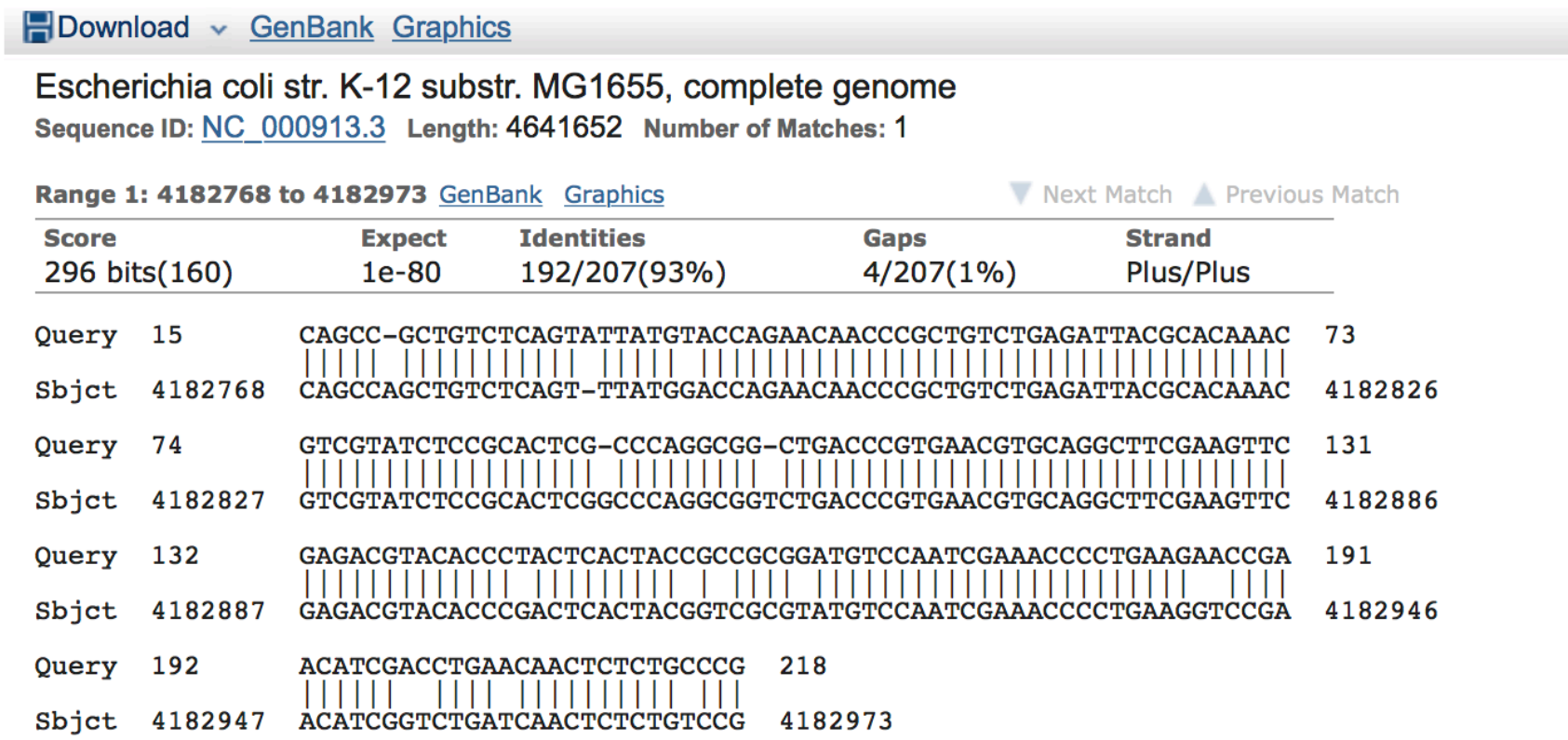
2019/15 BK II*

Sequence: 29415977

Samples:	16190	Quality: 0 - 9	
Bases:	249	10 - 19	
Average spacing:	66.0	20 - 29	
Average quality >=	10: 153, 20: 55, 30: 15	>= 30	

28.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182973 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
296 bits(160)	1e-80	192/207(93%)	4/207(1%)	Plus/Plus

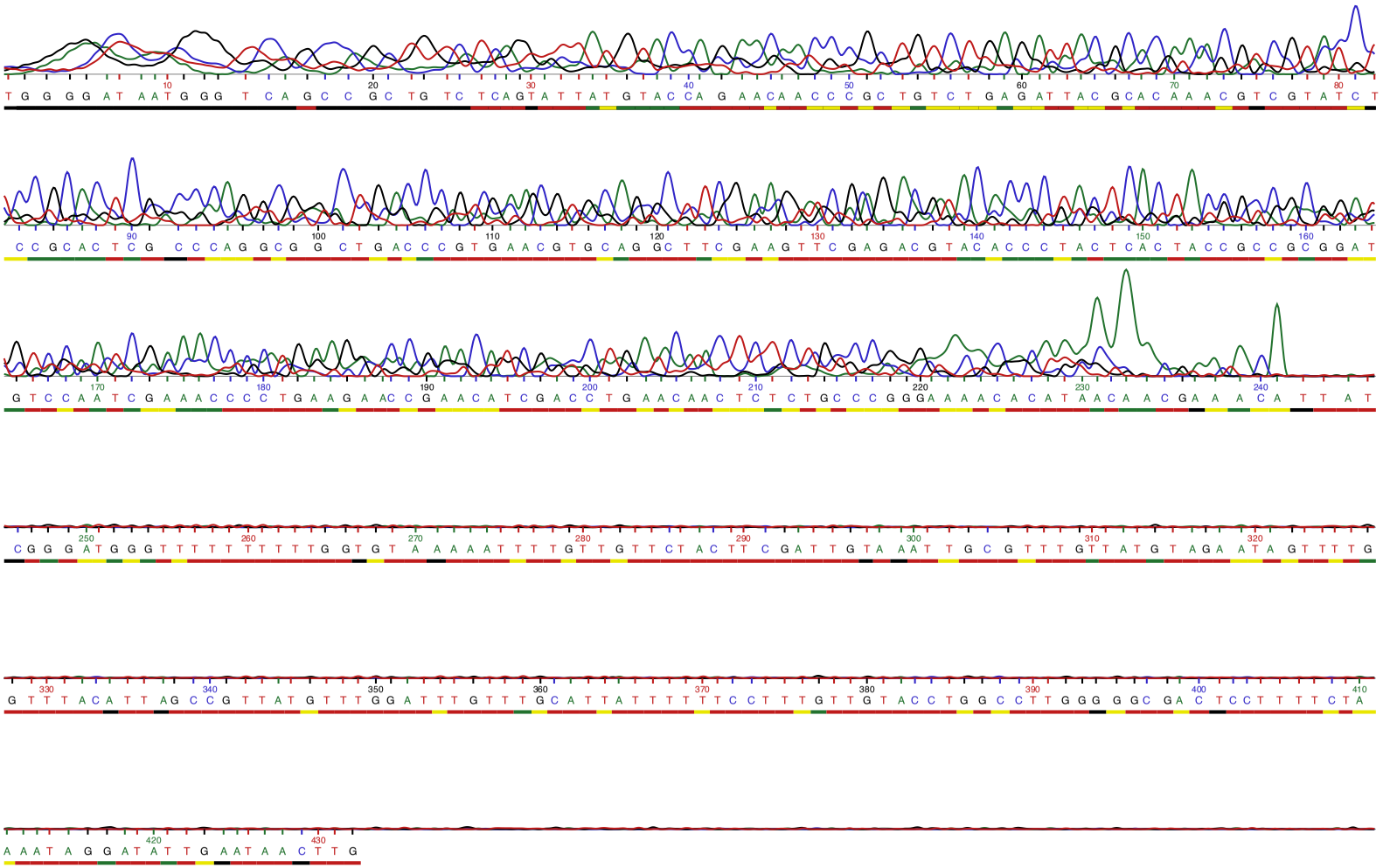
```
Query  15       CAGCC-GCTGTCTCAGTATTATGTACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  73
                |||||  |||||||||||  |||||  |||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGT-TTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  74       GTCGTATCTCCGCACTCG-CCCAGGCGG-CTGACCCGTGAACGTGCAGGCTTCGAAGTTC  131
                |||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  132      GAGACGTACACCCTACTCACTACCGCCGCGGATGTCCAATCGAAACCCCTGAAGAACCGA  191
                |||||||||||||  |||||||||  | ||||  ||||||||||||||||||||||  ||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  192      ACATCGACCTGAACAACTCTCTGCCCG  218
                ||||||   ||||  ||||||||||  |||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCG  4182973
```

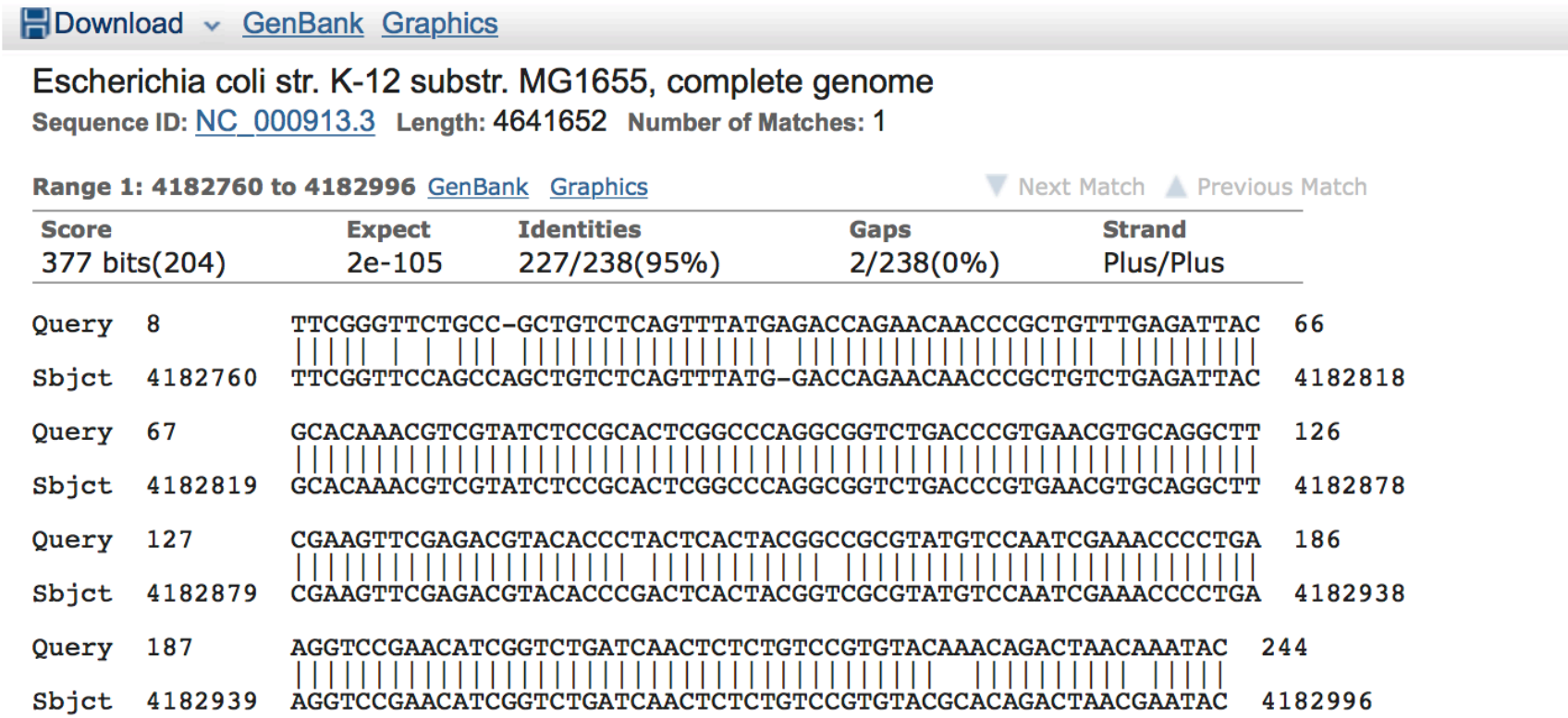
2019/18 AK I

Sequence: 29415243

Samples:	16300	Quality: 0 - 9	
Bases:	433	10 - 19	
Average spacing:	38.0	20 - 29	
Average quality >=	10: 254, 20: 87, 30: 50	>= 30	

22.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182760 to 4182996 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
377 bits(204)	2e-105	227/238(95%)	2/238(0%)	Plus/Plus

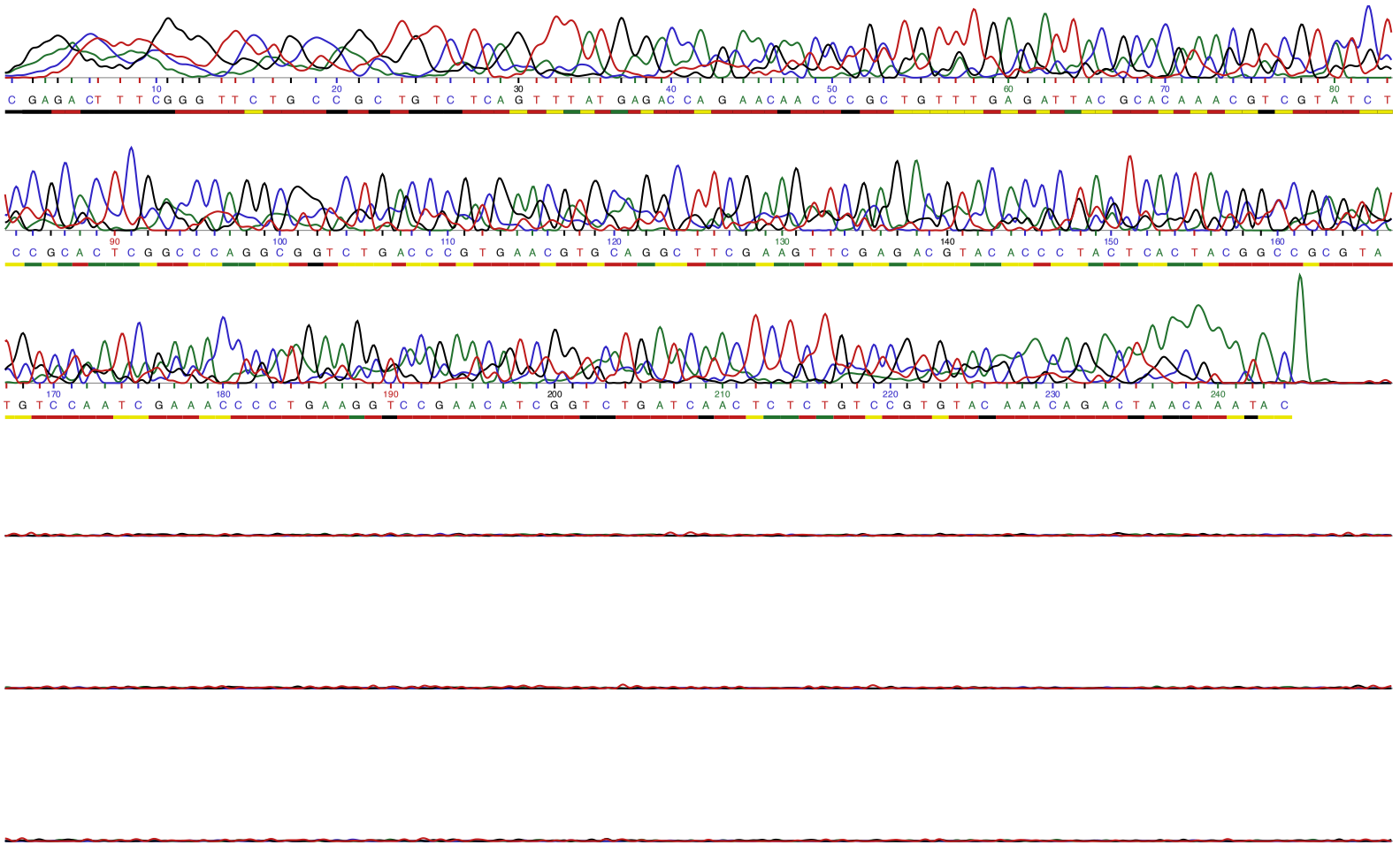
```
Query  8        TTCGGGTTCTGCC-GCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTTTGAGATTAC  66
               ||||| | | ||| ||||||||||||||| |||||||||||||||||||| |||||||||
Sbjct  4182760  TTCGGTTCCAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTAC  4182818

Query  67       GCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTT  126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182819  GCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTT  4182878

Query  127      CGAAGTTCGAGACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGA  186
               |||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Sbjct  4182879  CGAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGA  4182938

Query  187      AGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACAAACAGACTAACAAATAC  244
               |||||||||||||||||||||||||||||||||||||  |||||||||| |||||
Sbjct  4182939  AGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC  4182996
```

2019/18 CK I

Sequence: 29415274

Samples:	16303	Quality:	0 - 9
Bases:	245		10 - 19
Average spacing:	67.0		20 - 29
Average quality >=	10: 122, 20: 67, 30: 29		>= 30

22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182975 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
316 bits(171)	8e-87	196/208(94%)	2/208(0%)	Plus/Plus

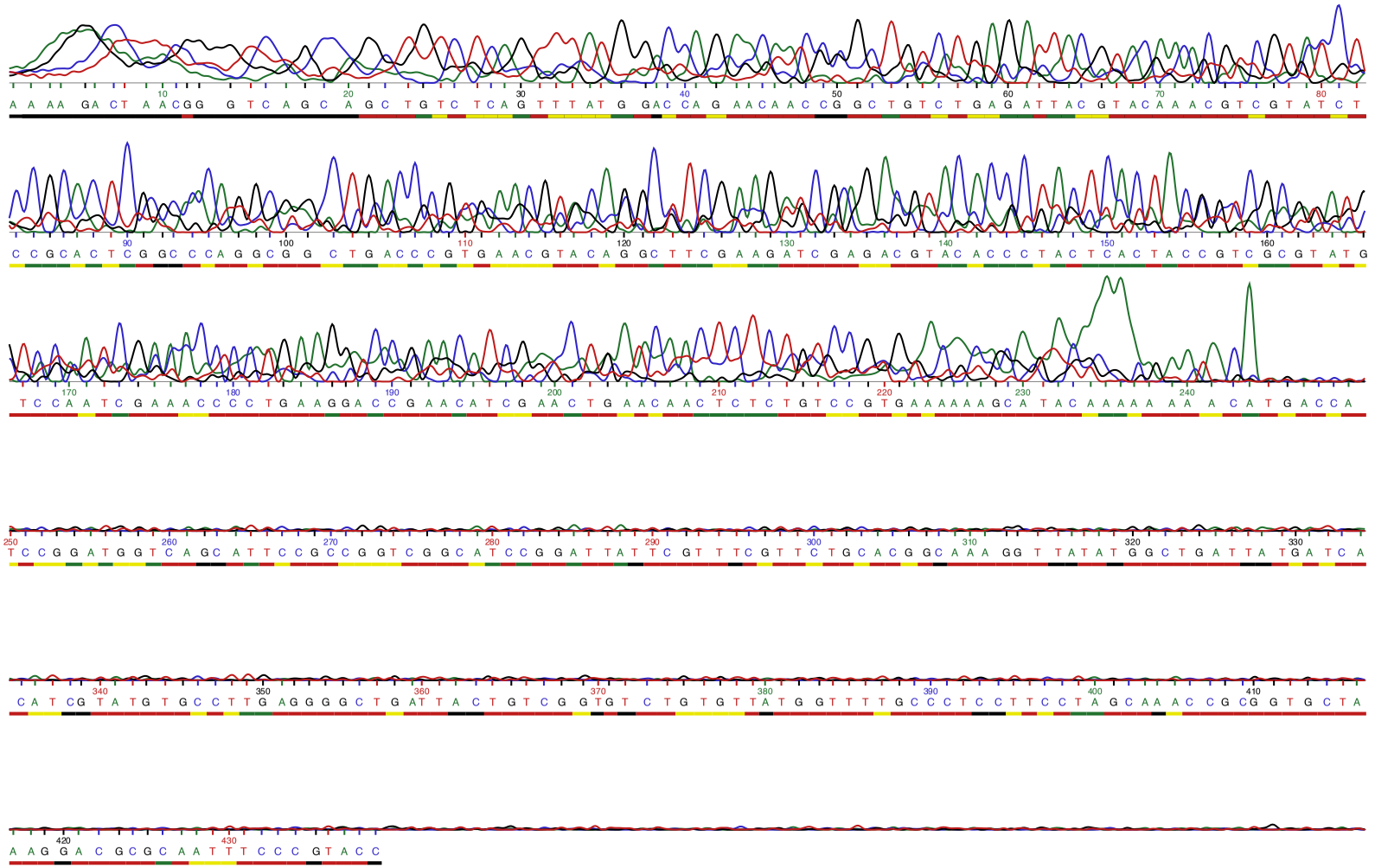
```
Query  16       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCGGCTGTCTGAGATTACGTACAAACG  74
                |||  |||||||||||||||||||||||||||||||| |||||||||||||||| ||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       TCGTATCTCCGCACTCGGCCCAGGCGG-CTGACCCGTGAACGTACAGGCTTCGAAGATCG  133
                |||||||||||||||||||||||||| ||||||||||||||| |||||||||||| |||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCTACTCACTACCGTCGCGTATGTCCAATCGAAACCCCTGAAGGACCGAA  193
                |||||||||||| ||||||||| ||||||||||||||||||||||||||||| |||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGAACTGAACAACTCTCTGTCCGTG  221
                |||||  |||| ||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTG  4182975
```

2019/20 BK I

Sequence: 29415595

Samples:	16300	Quality: 0 - 9	
Bases:	440	10 - 19	
Average spacing:	38.0	20 - 29	
Average quality >=	10: 239, 20: 92, 30: 61	>= 30	

22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
372 bits(201)	9e-104	211/216(98%)	0/216(0%)	Plus/Plus

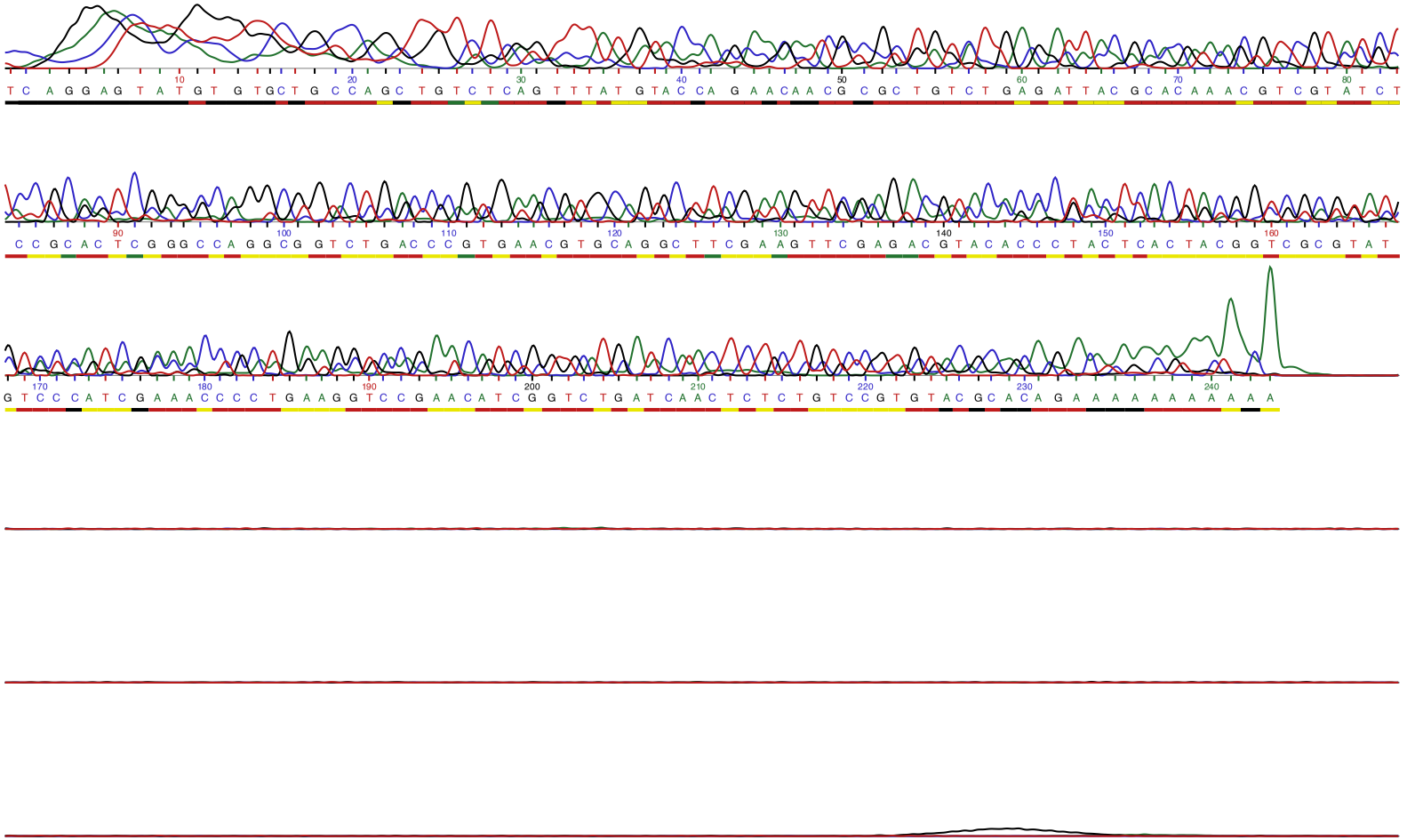
```
Query  18       GCCAGCTGTCTCAGTTTATGTACCAGAACAACGCGCTGTCTGAGATTACGCACAAACGTC  77
                ||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  78       GTATCTCCGCACTCGGGCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  137
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  138      ACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAACA  197
                |||||||||| |||||||||||||||||||||| |||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  198      TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGa  233
                ||||||||||||||||||||||||||||||||||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

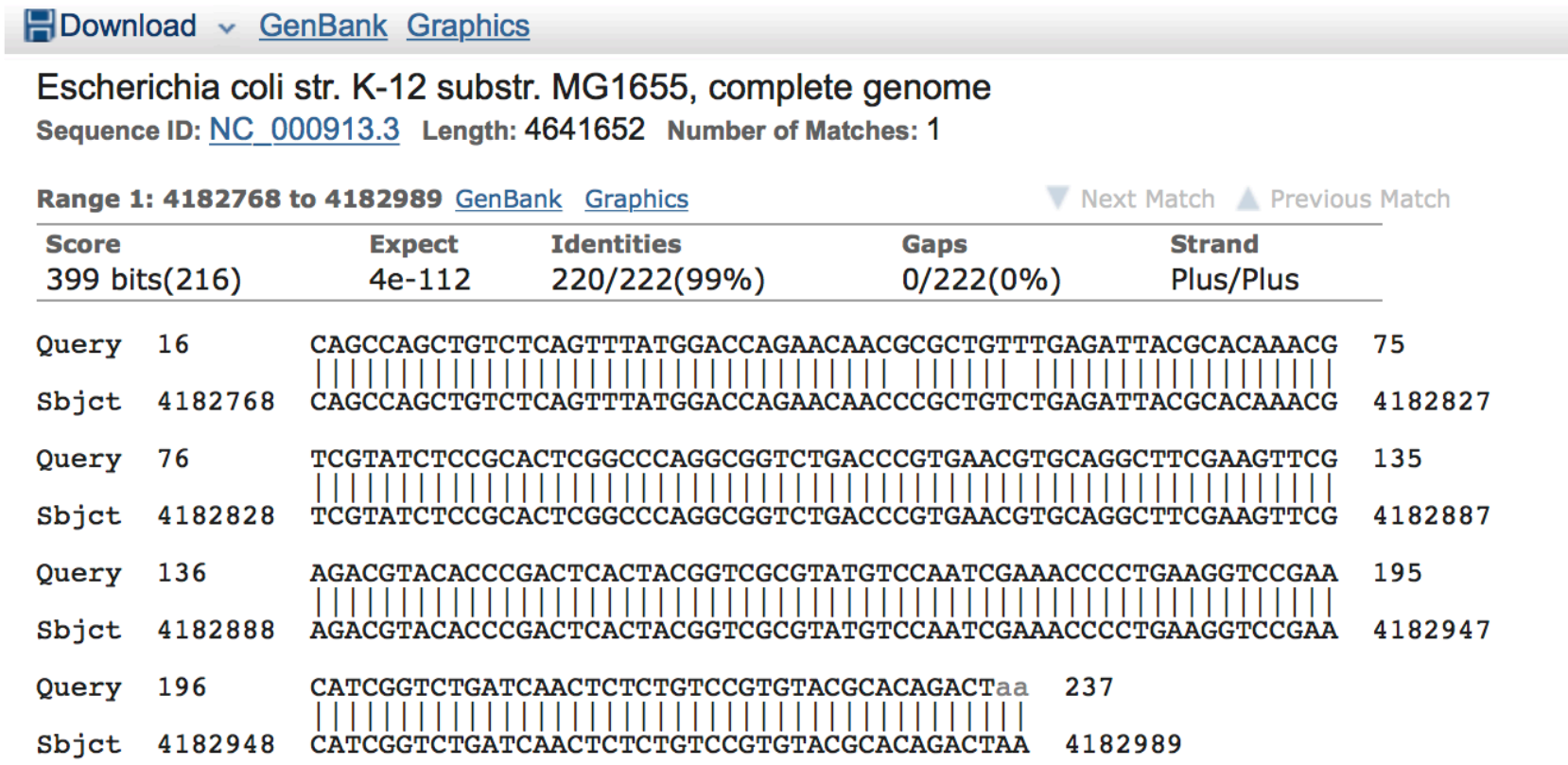
2019/23 CK I

Sequence: 29416530

Samples:	16300	Quality:	0 - 9
Bases:	244		10 - 19
Average spacing:	67.0		20 - 29
Average quality >=	10: 123, 20: 79, 30: 9		>= 30

20.06.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182989 GenBank Graphics Next Match Previous Match

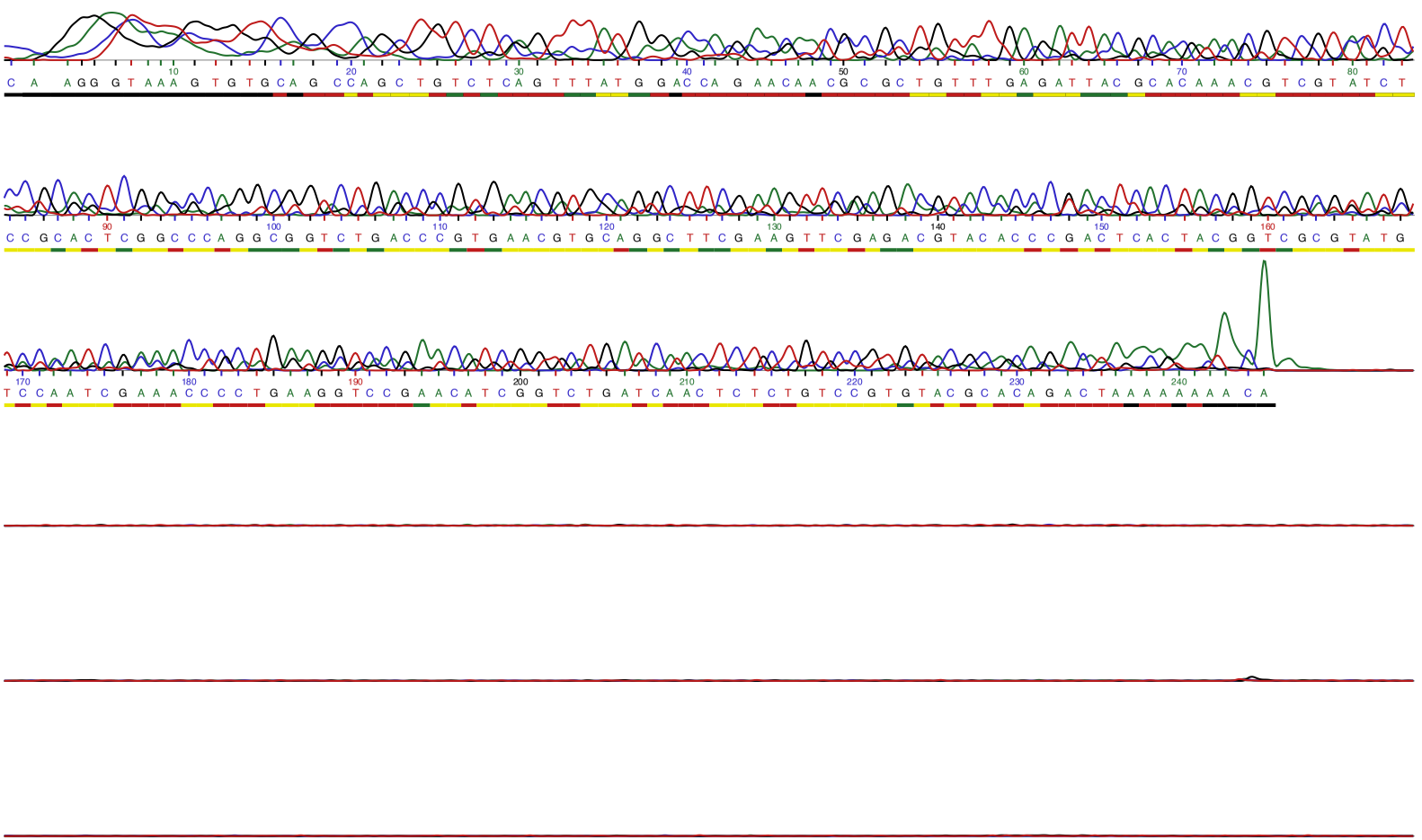
Score	Expect	Identities	Gaps	Strand
399 bits(216)	4e-112	220/222(99%)	0/222(0%)	Plus/Plus

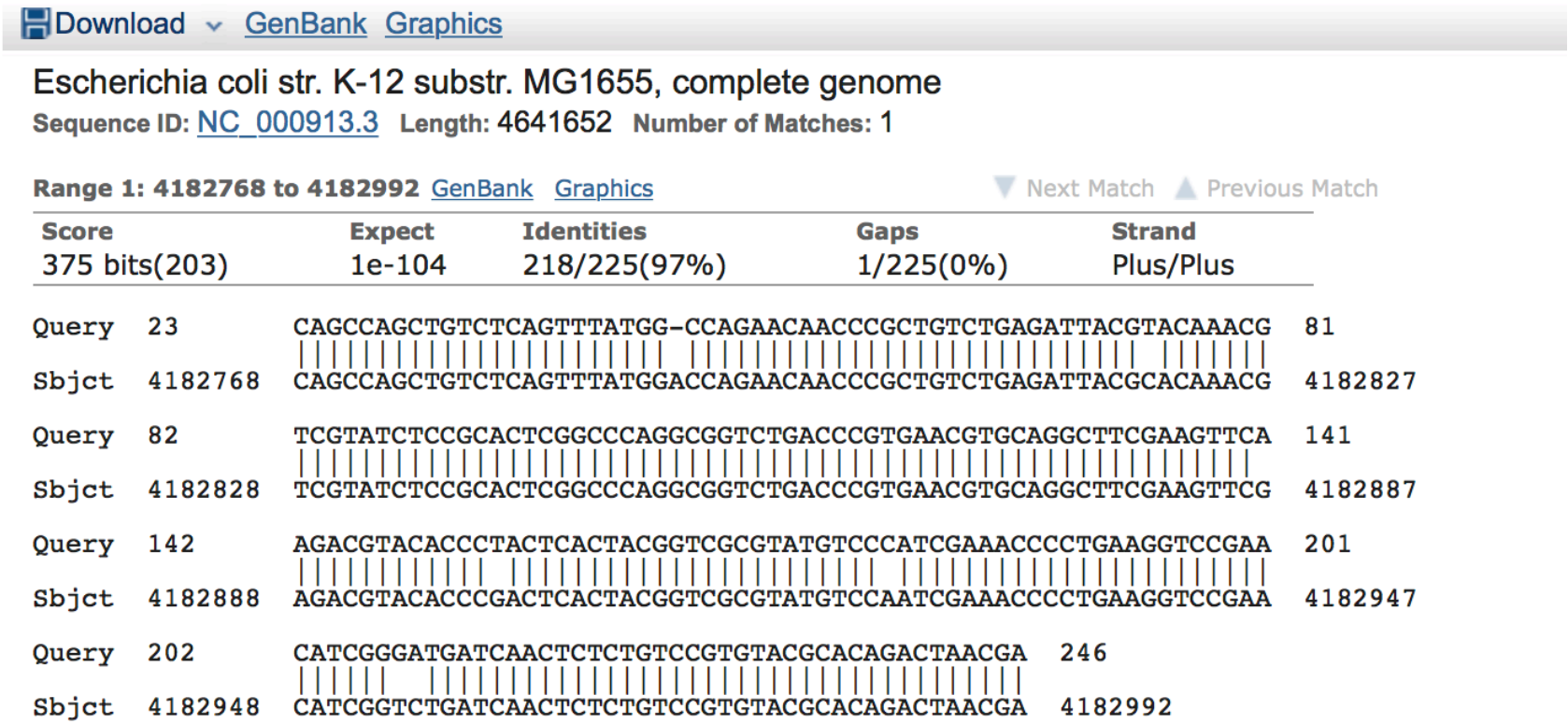
```
Query  16       CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACGCGCTGTTTGAGATTACGCACAAACG  75
                ||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  76       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  136      AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  196      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  237
                ||||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/23 CK III

Sequence: 29416691

Samples:	16299	Quality:	0 - 9
Bases:	246		10 - 19
Average spacing:	67.0		20 - 29
Average quality >=	10: 88, 20: 103, 30: 30		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182992 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
375 bits(203)	1e-104	218/225(97%)	1/225(0%)	Plus/Plus

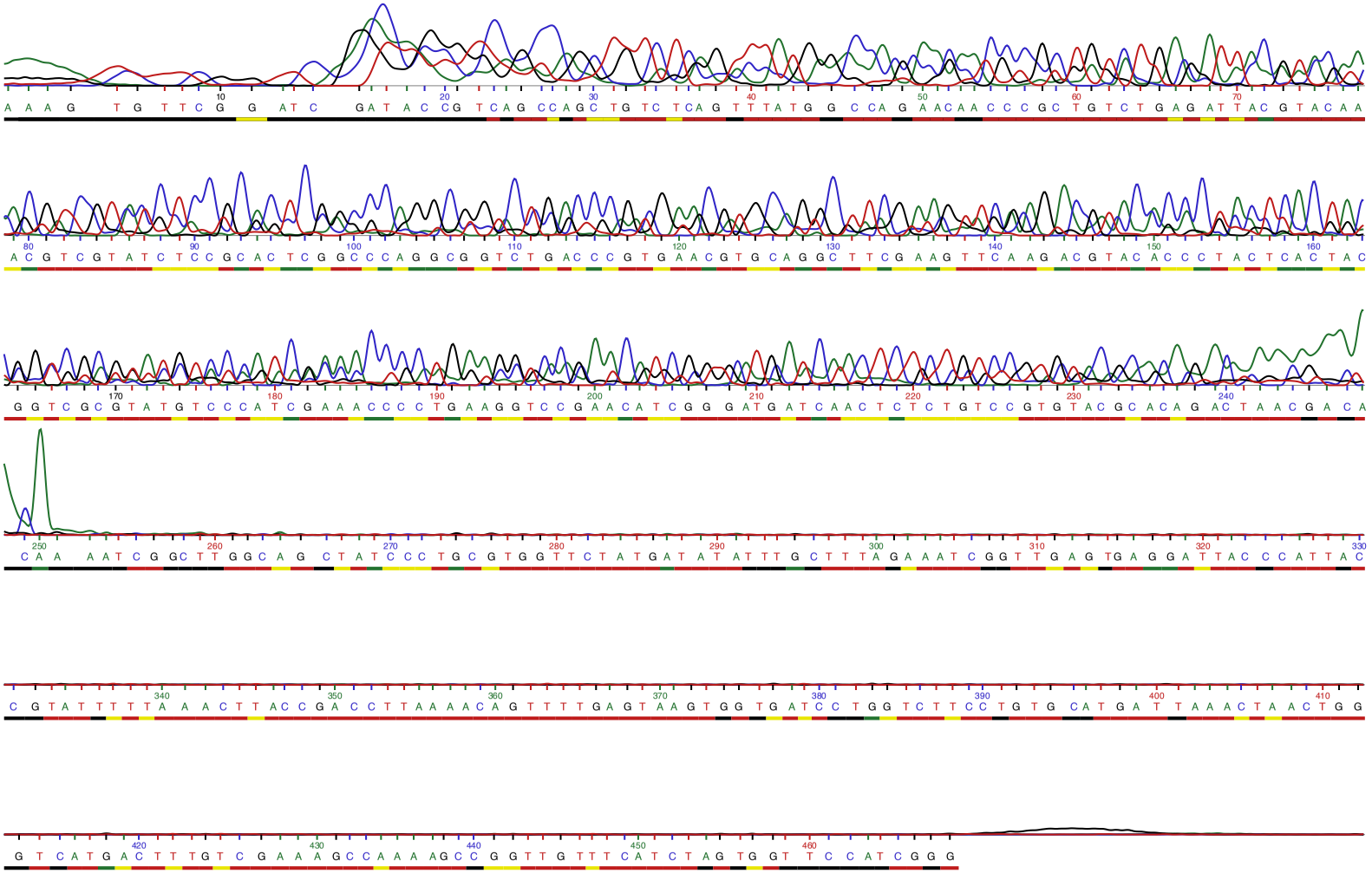
```
Query  23       CAGCCAGCTGTCTCAGTTTATGG-CCAGAACAACCCGCTGTCTGAGATTACGTACAAACG  81
                |||||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  82       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCA  141
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  142      AGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAA  201
                |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  202      CATCGGGATGATCAACTCTCTGTCCGTGTACGCACAGACTAACGA  246
                ||||||  ||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGA  4182992
```

2019/24 CK I

Sequence: 29416776

Samples:	16302	Quality: 0 - 9	
Bases:	469	10 - 19	
Average spacing:	35.0	20 - 29	
Average quality >=	10: 263, 20: 93, 30: 39	>= 30	

Page: 1 / 3
20.06.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
368 bits(199)	1e-102	215/222(97%)	3/222(1%)	Plus/Plus

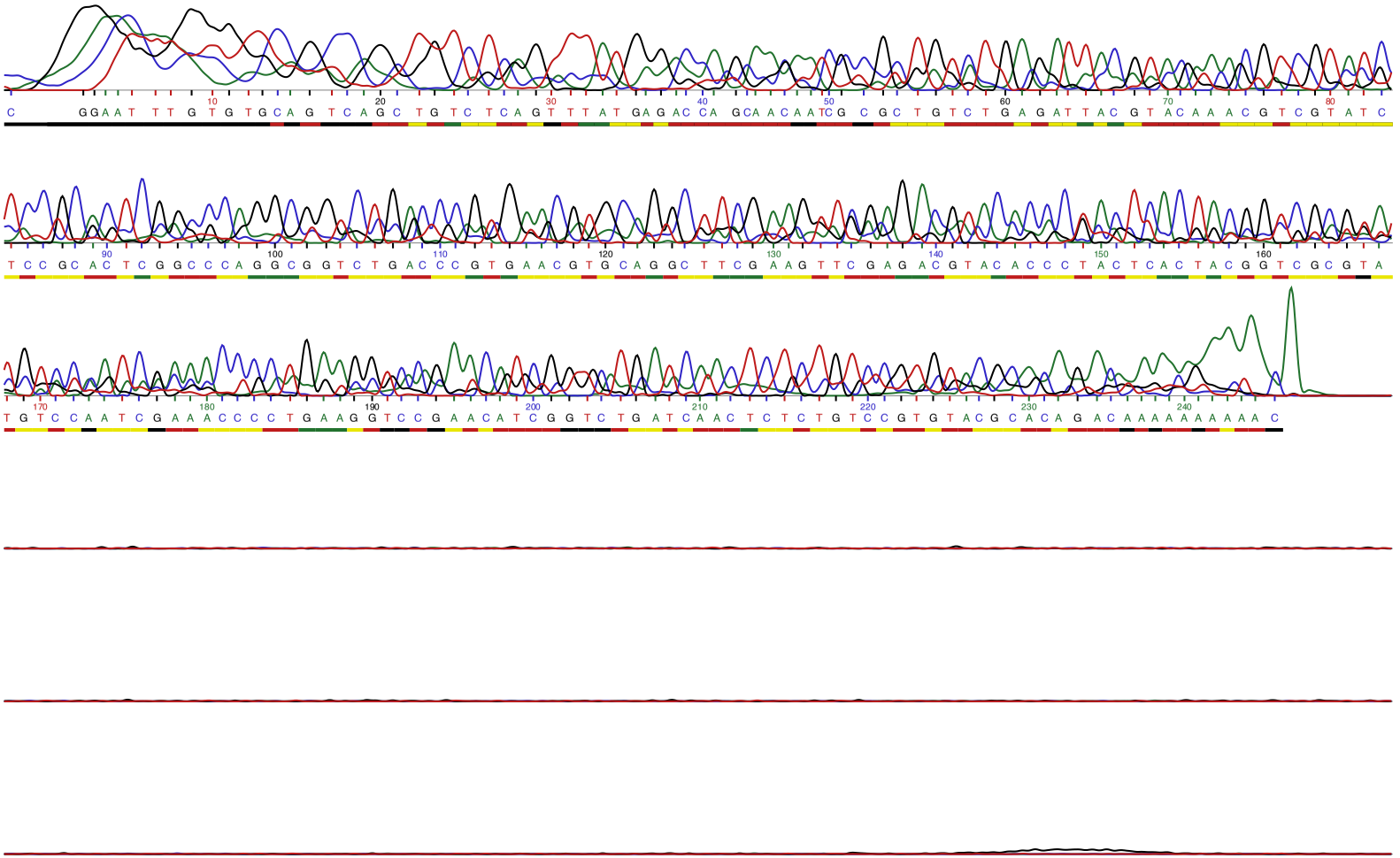
```
Query  14       CAGTCAGCTGTCTCAGTTTATGAGACCAGCAACAATCGCGCTGTCTGAGATTACGTACAA  73
                |||  ||||||||||||||||||  ||||||  |||||  |  ||||||||||||||||  ||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAG-AACAA-CCCGCTGTCTGAGATTACGCACAA  4182824

Query  74       ACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGT  133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182825  ACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGT  4182884

Query  134      TCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCC  193
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182885  TCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCC  4182944

Query  194      GAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  235
                ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182945  GAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

2019/24 DK I

Sequence: 29416813

Samples:	16303	Quality: 0 - 9	
Bases:	247	10 - 19	
Average spacing:	67.0	20 - 29	
Average quality >=	10: 94, 20: 93, 30: 25	>= 30	

20.06.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182972 GenBank Graphics Next Match Previous Match

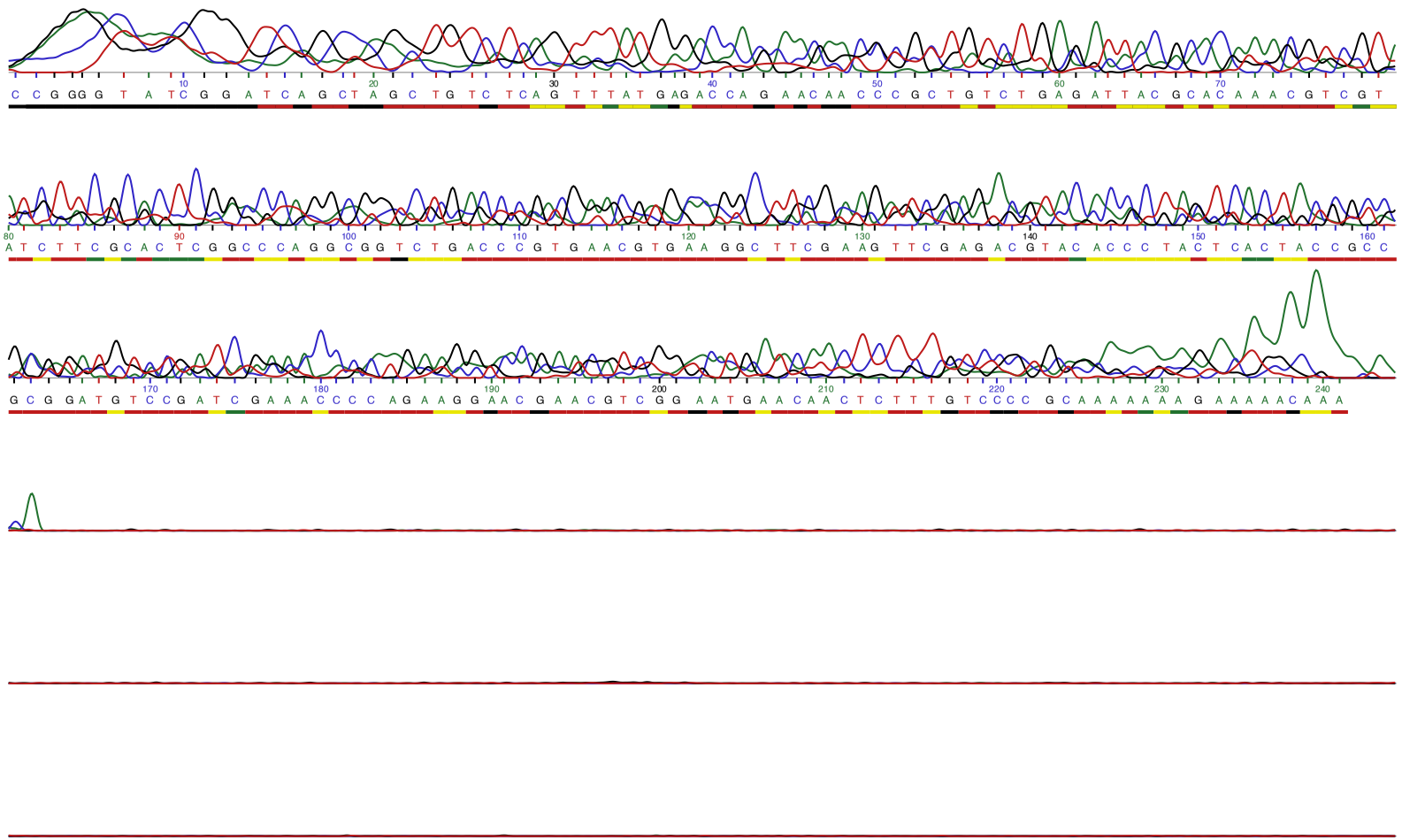
Score	Expect	Identities	Gaps	Strand
287 bits(155)	3e-78	194/213(91%)	2/213(0%)	Plus/Plus

```
Query  9        TCGGAT-CAGCTAGCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTCTGAGATTACG  67
               |||| | |||| ||||||||||||||||| |||||||||||||||||||||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACG  4182819

Query  68       CACAAACGTCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGAAGGCTTC  127
               ||||||||||||||| |||||||||||||||||||||||||||||||||||| |||||||
Sbjct  4182820  CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  4182879

Query  128      GAAGTTCGAGACGTACACCCTACTCACTACCGCCGCGGATGTCCGATCGAAACCCCAGAA  187
               |||||||||||||||||||| ||||||||| | |||| |||||| |||||||||||| |||
Sbjct  4182880  GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  4182939

Query  188      GGAACGAACGTCGGAATGAACAACTCTTTGTCC  220
               ||  ||||| |||| ||| |||||||| |||||
Sbjct  4182940  GGTCCGAACATCGGTCTGATCAACTCTCTGTCC  4182972
```

2016/1 AK I

Sequence: 50401659

Samples:	15232	Quality: 0 - 9	
Bases:	242	10 - 19	
Average spacing:	63.0	20 - 29	
Average quality >=	10: 136, 20: 58, 30: 14	>= 30	



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182953 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
298 bits(161)	2e-81	178/186(96%)	1/186(0%)	Plus/Plus

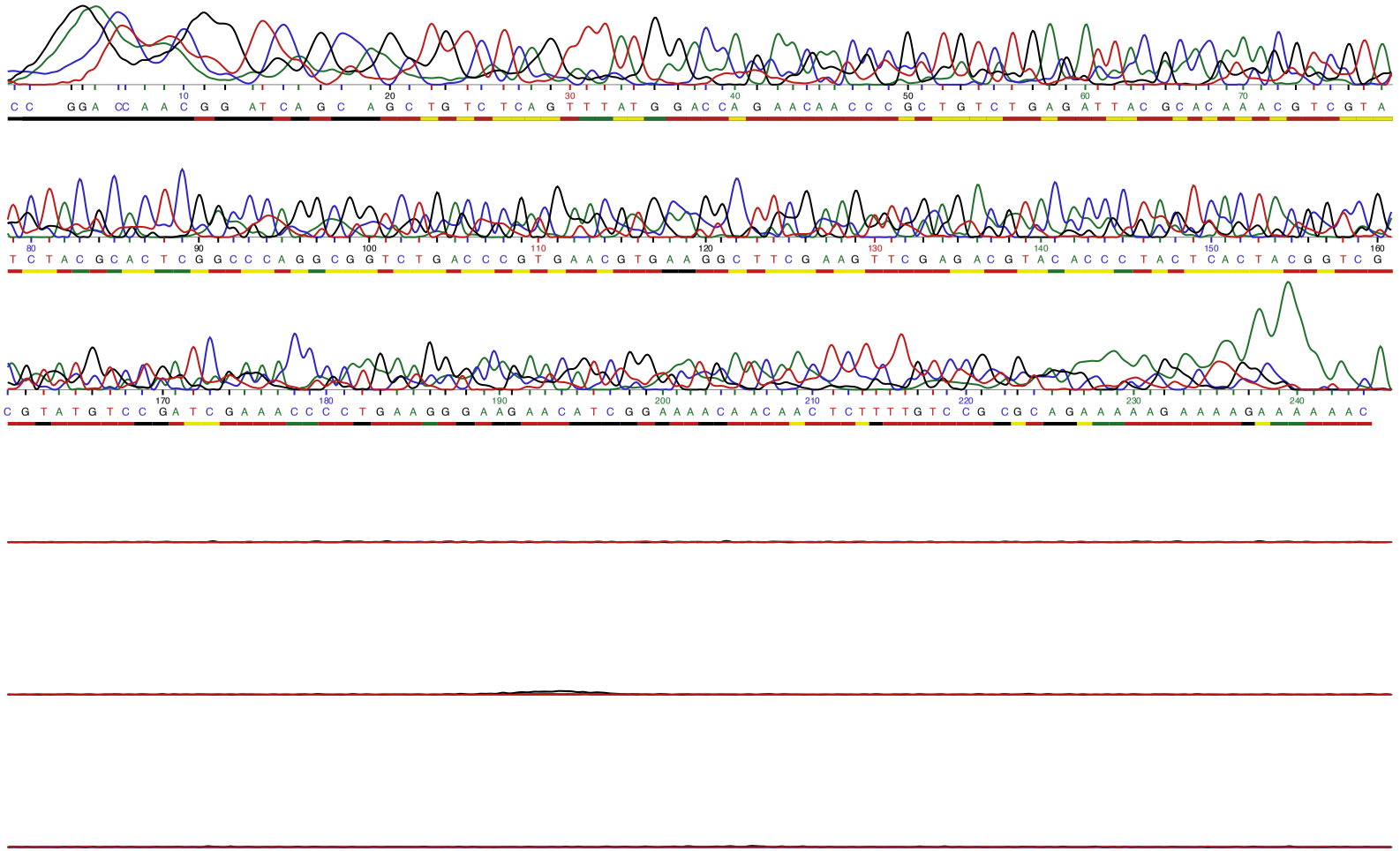
```
Query  15       CAG-CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  73
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  74       TCGTATCTACGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGAAGGCTTCGAAGTTCG  133
               |||||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCTACTCACTACGGTCGCGTATGTCCGATCGAAACCCCTGAAGGGAAGAA  193
               |||||||||||| |||||||||||||||||||| |||||||||||||||||   |||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGG  199
               ||||||
Sbjct  4182948  CATCGG  4182953
```

2016/1 AK II

Sequence: 50401642

Samples:	15247	Quality: 0 - 9	
Bases:	245	10 - 19	
Average spacing:	63.0	20 - 29	
Average quality >=	10: 120, 20: 71, 30: 17	>= 30	

Page: 1 / 3
29.03.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182761 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
368 bits(199)	1e-102	219/228(96%)	3/228(1%)	Plus/Plus

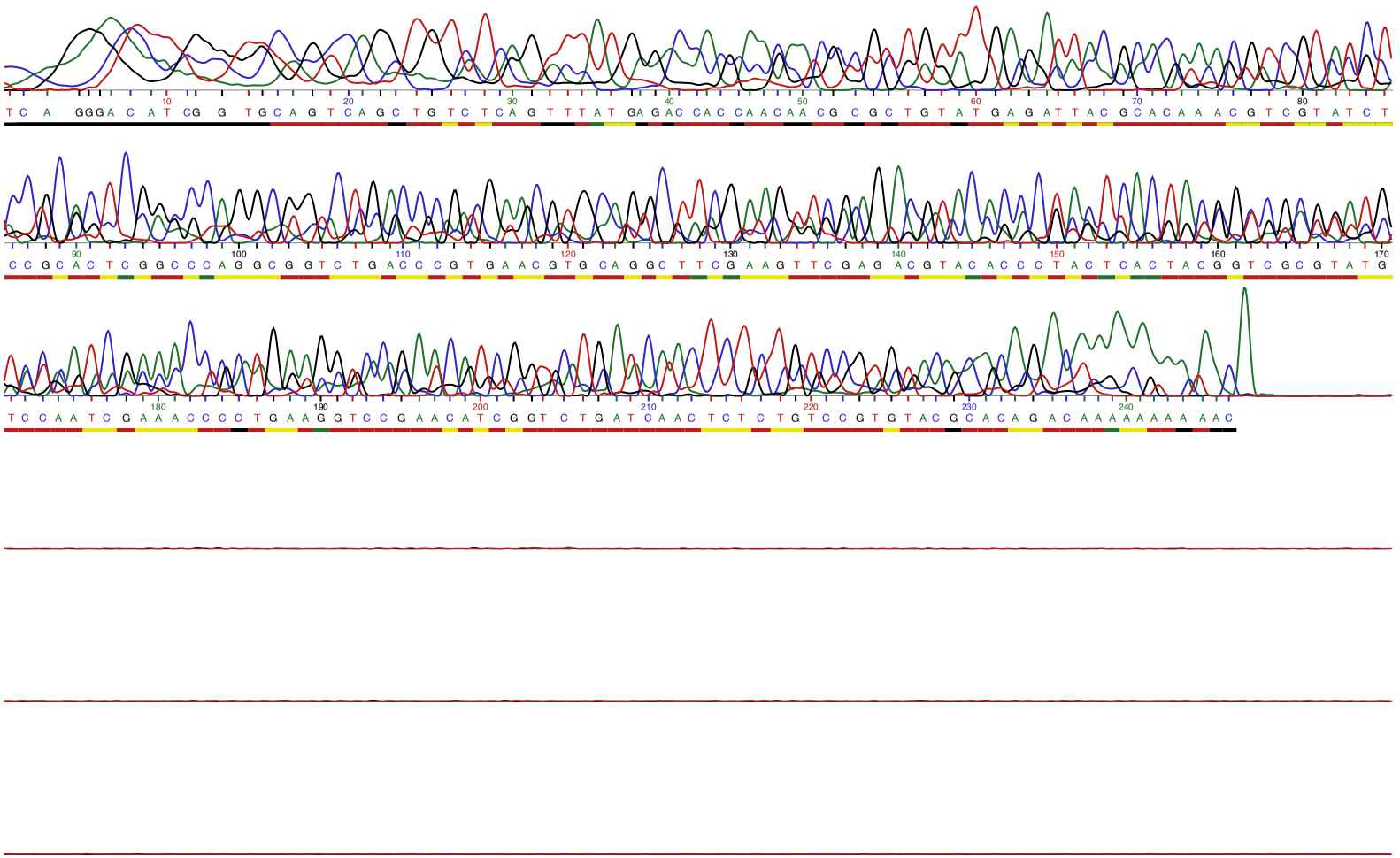
```
Query  10       TCGG-TGCAGTCAGCTGTCTCAGTTTATGAGACCACCAACAACGCGCTGTATGAGATTAC  68
               ||||  |  |||  ||||||||||||||||||||  |||||   |||||| ||||||  ||||||||||
Sbjct  4182761  TCGGTTCCAGCCAGCTGTCTCAGTTTATG-GACCA-GAACAACCCGCTGTCTGAGATTAC  4182818

Query  69       GCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTT  128
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182819  GCACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTT  4182878

Query  129      CGAAGTTCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGA  188
               |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4182879  CGAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGA  4182938

Query  189      AGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  236
               ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182939  AGGTCCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

2016/2 AK III

Sequence: 29412563

Samples:	15257	Quality:	0 - 9
Bases:	248		10 - 19
Average spacing:	62.0		20 - 29
Average quality >=	10: 133, 20: 73, 30: 11		>= 30

16.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182990 GenBank Graphics Next Match Previous Match

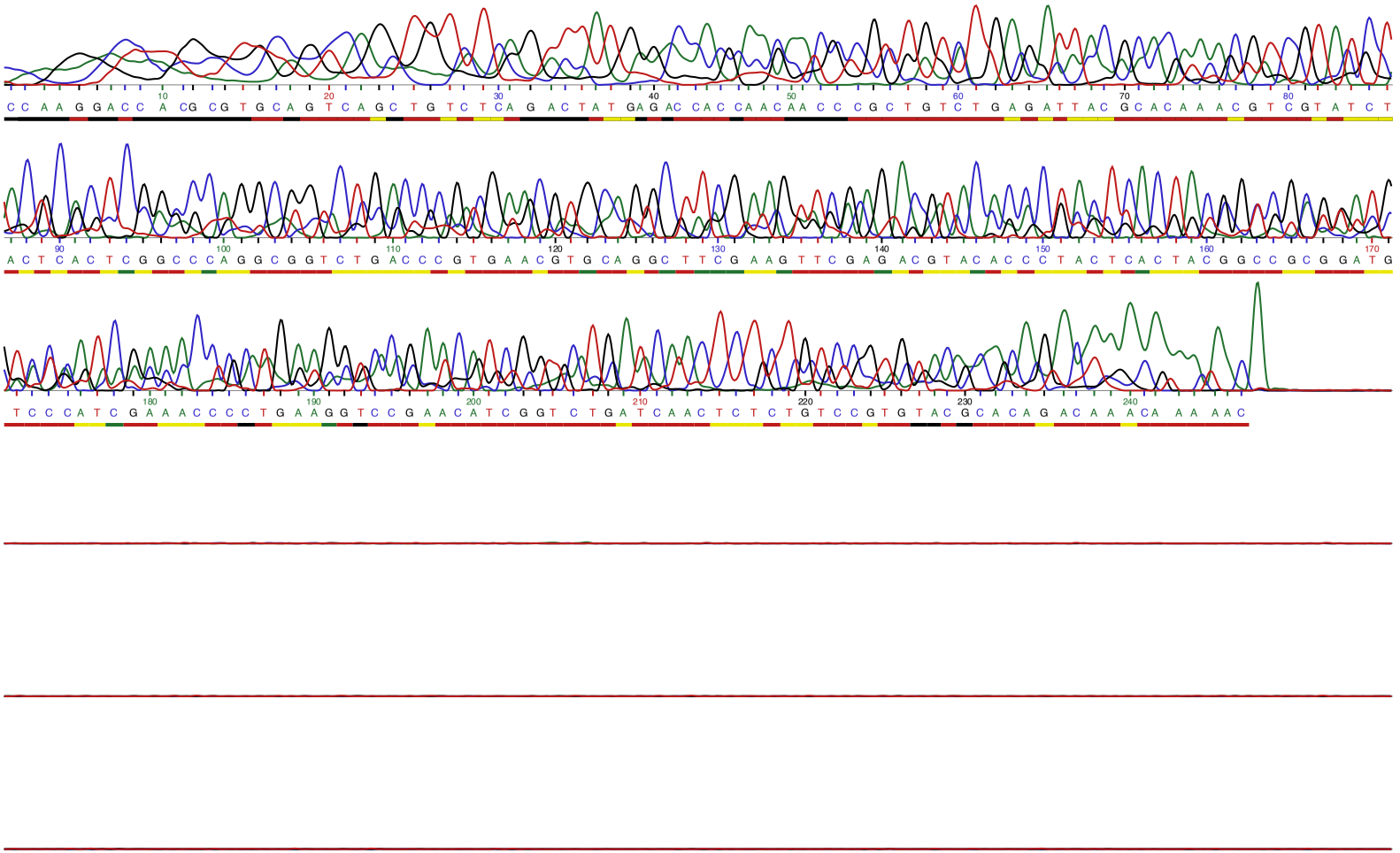
Score	Expect	Identities	Gaps	Strand
342 bits(185)	7e-95	212/225(94%)	2/225(0%)	Plus/Plus

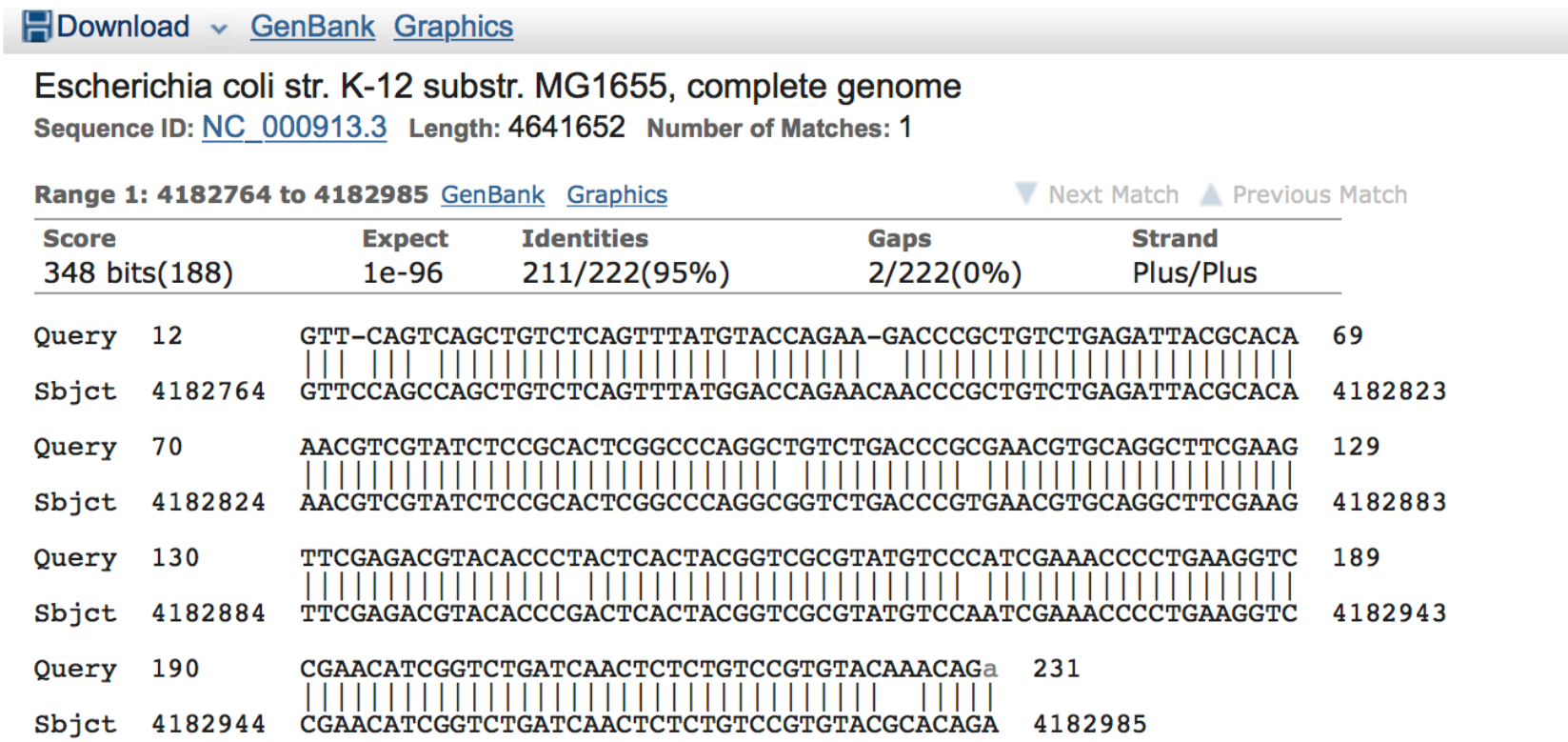
```
Query  17       CAGTCAGCTGTCTCAGACTATGAGACCACCAACAACCCGCTGTCTGAGATTACGCACAAA  76
                |||  |||||||||||||   ||||  |||||  ||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCA-GAACAACCCGCTGTCTGAGATTACGCACAAA  4182825

Query  77       CGTCGTATCTACTCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  136
                ||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182826  CGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  4182885

Query  137      CGAGACGTACACCCTACTCACTACGGCCGCGGATGTCCCATCGAAACCCCTGAAGGTCCG  196
                |||||||||||||| |||||||||||  ||||  ||||||  ||||||||||||||||||||
Sbjct  4182886  CGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCG  4182945

Query  197      AACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACAAAC  241
                |||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  4182946  AACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAAC  4182990
```

2016/2 AK IV

Sequence: 29412747

Samples:	15211	Quality:	0 - 9
Bases:	248		10 - 19
Average spacing:	62.0		20 - 29
Average quality >=	10: 135, 20: 69, 30: 13		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182764 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
348 bits(188)	1e-96	211/222(95%)	2/222(0%)	Plus/Plus

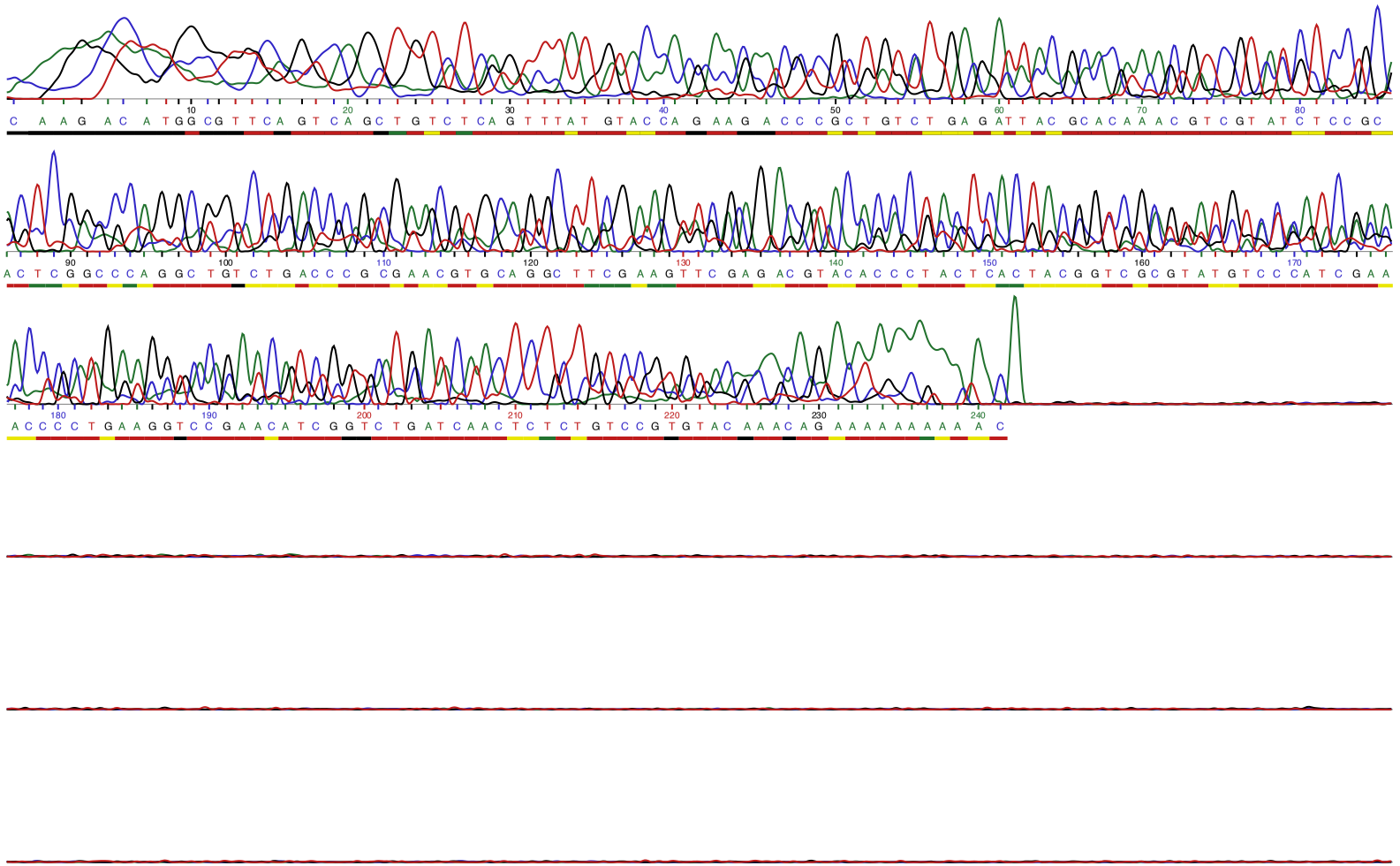
```
Query  12       GTT-CAGTCAGCTGTCTCAGTTTATGTACCAGAA-GACCCGCTGTCTGAGATTACGCACA  69
               ||| ||| ||||||||||||||||||| ||||||| ||||||||||||||||||||||||
Sbjct  4182764  GTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACA  4182823

Query  70       AACGTCGTATCTCCGCACTCGGCCCAGGCTGTCTGACCCGCGAACGTGCAGGCTTCGAAG  129
               ||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Sbjct  4182824  AACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAG  4182883

Query  130      TTCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTC  189
               |||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Sbjct  4182884  TTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTC  4182943

Query  190      CGAACATCGGTCTGATCAACTCTCTGTCCGTGTACAAACAGa  231
               |||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  4182944  CGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2016/2 AK V

Sequence: 29412785

Samples:	15980	Quality: 0 - 9	
Bases:	242	10 - 19	
Average spacing:	67.0	20 - 29	
Average quality >=	10: 149, 20: 55, 30: 14	>= 30	

Page: 1 / 3
16.05.2019



2016/3 AK III

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182773 to 4182932 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
248 bits(134)	2e-66	153/162(94%)	2/162(1%)	Plus/Plus

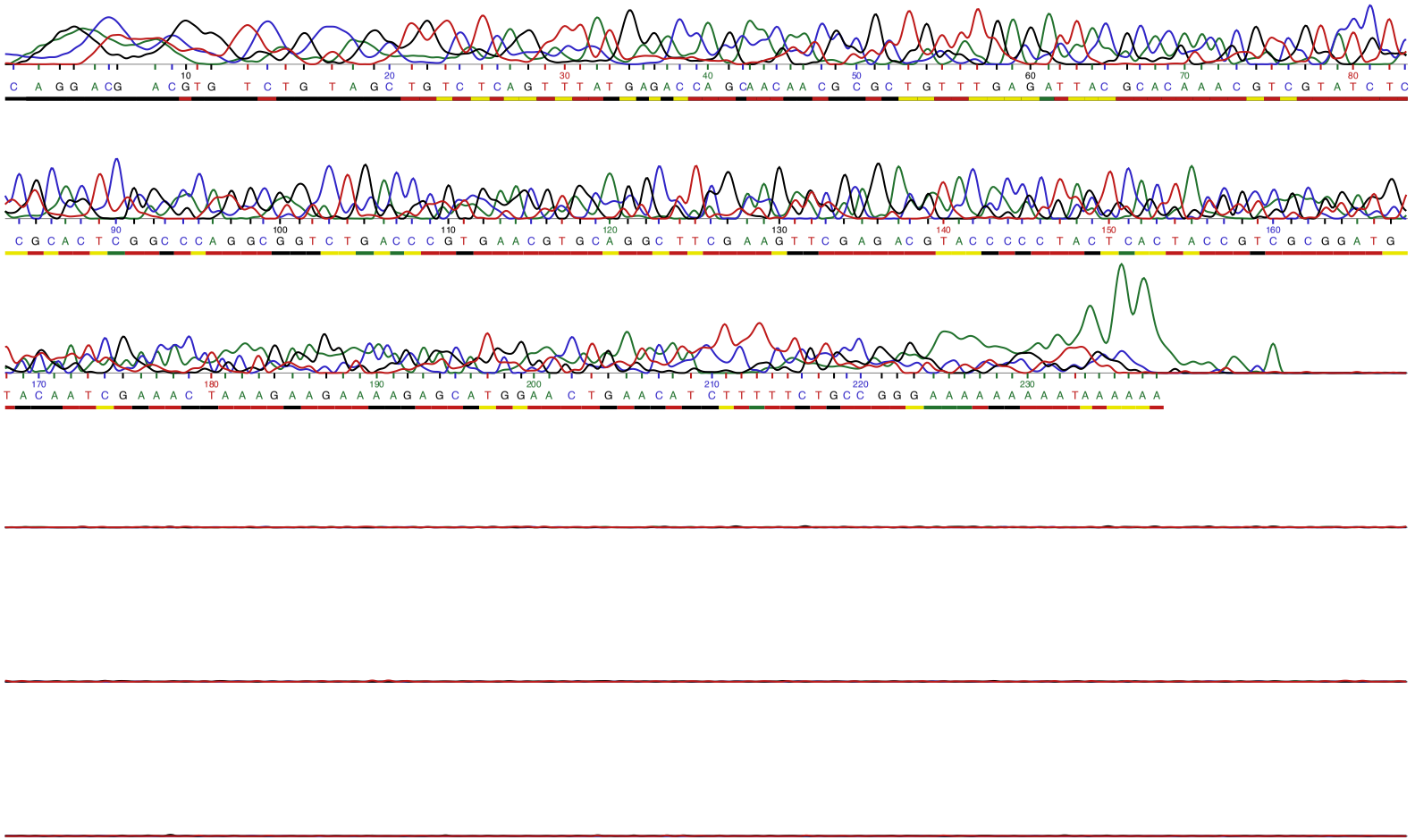
```
Query  18       AGCTGTCTCAGTTTATGAGACCAGCAACAACGCGCTGTTTGAGATTACGCACAAACGTCG  77
                ||||||||||||||||| ||||||| ||||||| ||||||| ||||||||||||||||||||
Sbjct  4182773  AGCTGTCTCAGTTTATG-GACCAG-AACAACCCGCTGTCTGAGATTACGCACAAACGTCG  4182830

Query  78       TATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGA  137
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182831  TATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGA  4182890

Query  138      CGTACCCCCTACTCACTACCGTCGCGGATGTACAATCGAAAC  179
                ||||| ||| |||||||||| |||||| |||| ||||||||||
Sbjct  4182891  CGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAAC  4182932
```

2016/3 AK III

Sequence: 29414048

Samples:	16300	Quality: 0 - 9	
Bases:	240	10 - 19	
Average spacing:	68.0	20 - 29	
Average quality >=	10: 124, 20: 48, 30: 9	>= 30	

Page: 1 / 3
22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182977 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
329 bits(178)	9e-91	199/209(95%)	2/209(0%)	Plus/Plus

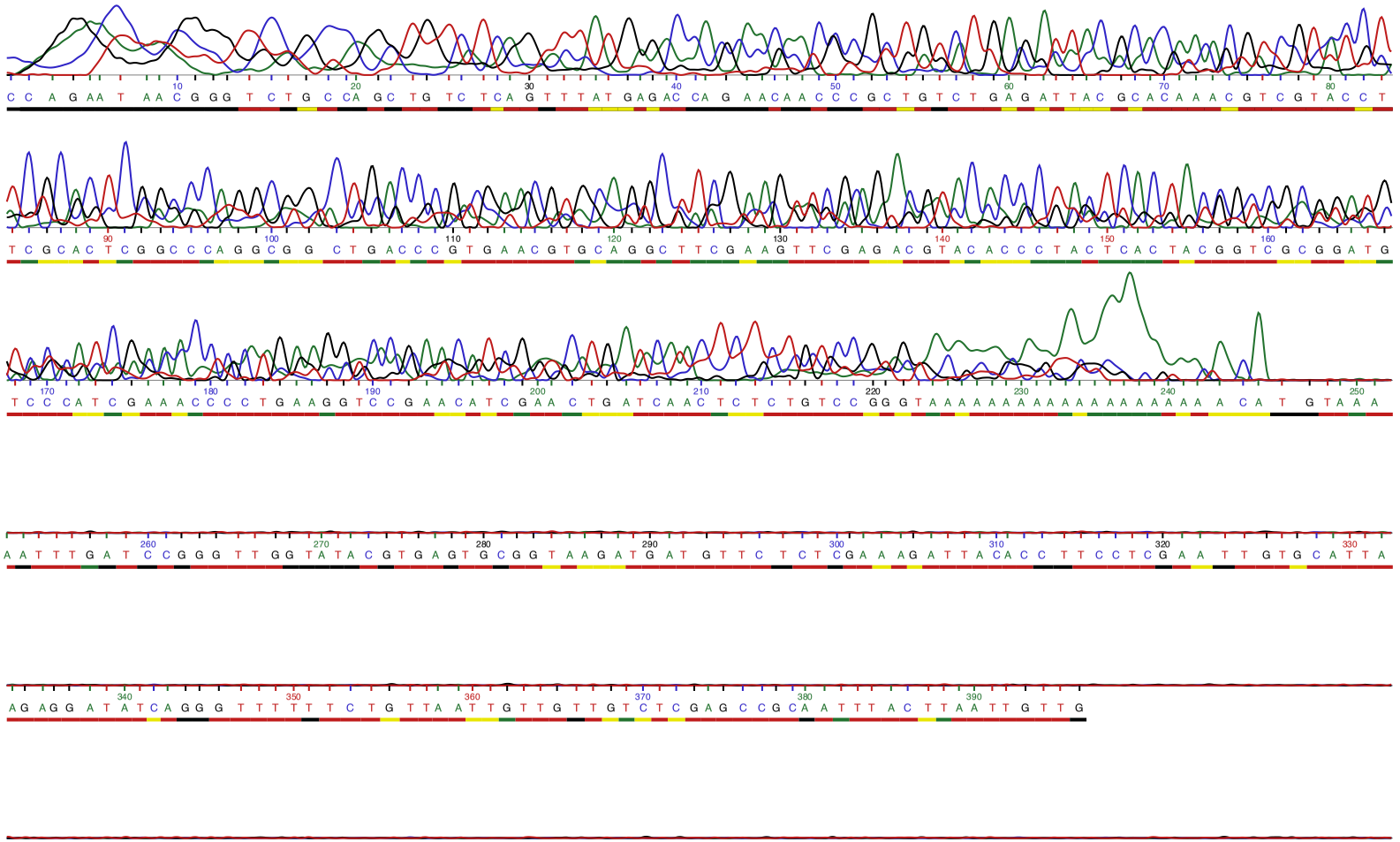
```
Query  17       GCCAGCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  76
                ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  4182828

Query  77       CGTACCTTCGCACTCGGCCCAGGCGG-CTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  135
                ||||  ||  |||||||||||||||| |||||||||||||||||||||||||||||||
Sbjct  4182829  CGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  4182888

Query  136      GACGTACACCCTACTCACTACGGTCGCGGATGTCCCATCGAAACCCCTGAAGGTCCGAAC  195
                |||||||||||  |||||||||||||||  |||||||  |||||||||||||||||||||||
Sbjct  4182889  GACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAAC  4182948

Query  196      ATCGAACTGATCAACTCTCTGTCCGGGTa  224
                ||||  |||||||||||||||||||| |||
Sbjct  4182949  ATCGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

2016/3 AK IV

Sequence: 29414123

Samples:	16302	Quality: 0 - 9	
Bases:	397	10 - 19	
Average spacing:	42.0	20 - 29	
Average quality >=	10: 223, 20: 75, 30: 43	>= 30	

22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182977 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
327 bits(177)	2e-90	198/208(95%)	2/208(0%)	Plus/Plus

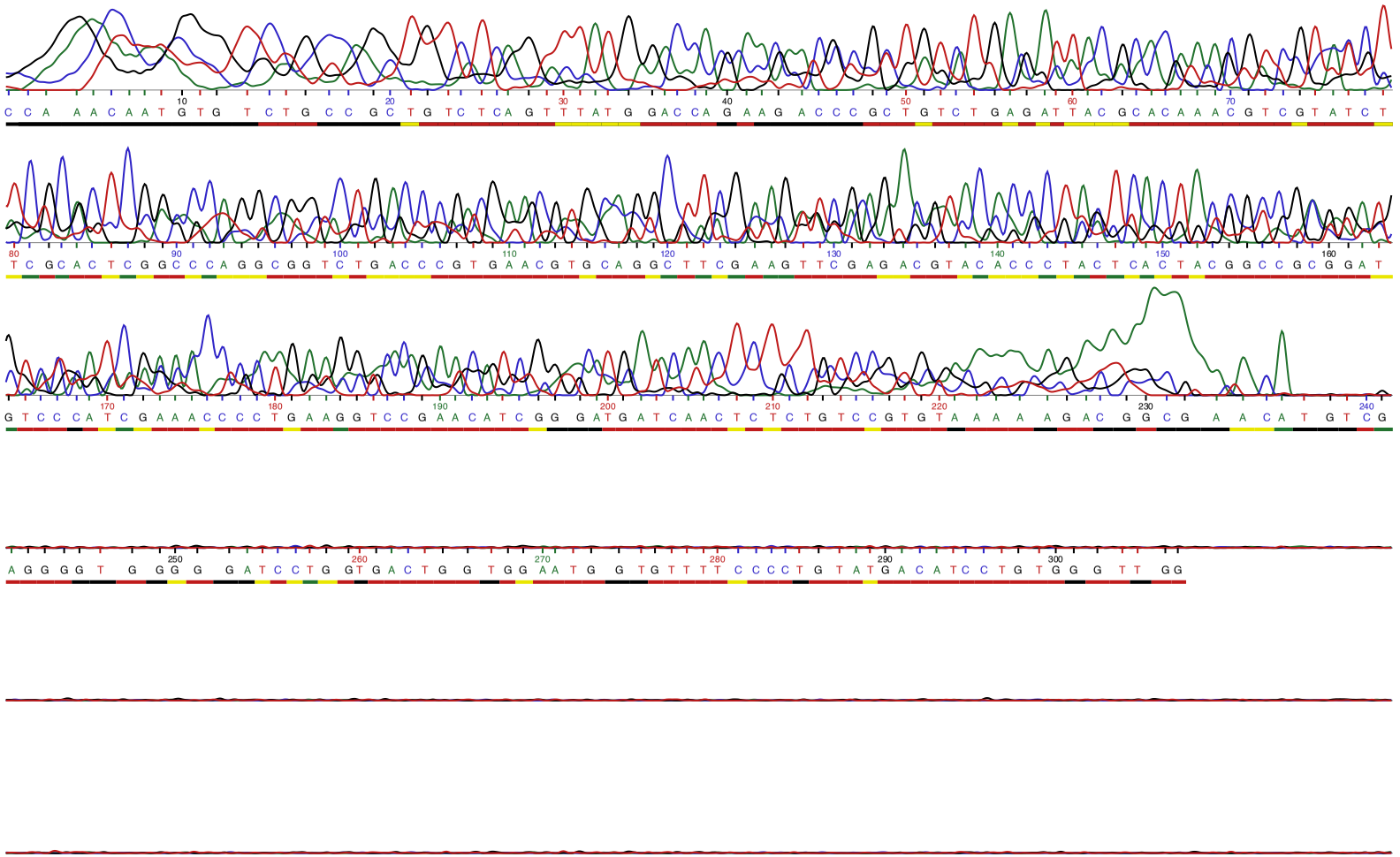
```
Query  16       GCC-GCTGTCTCAGTTTATGGACCAGAA-GACCCGCTGTCTGAGATTACGCACAAACGTC  73
                |||  ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  74       GTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  133
                ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  134      ACGTACACCCTACTCACTACGGCCGCGGATGTCCCATCGAAACCCCTGAAGGTCCGAACA  193
                ||||||||||  ||||||||||  ||||  ||||||  ||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  194      TCGGGATGATCAACTCTCTGTCCGTGTA  221
                ||||  ||||||||||||||||||||||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

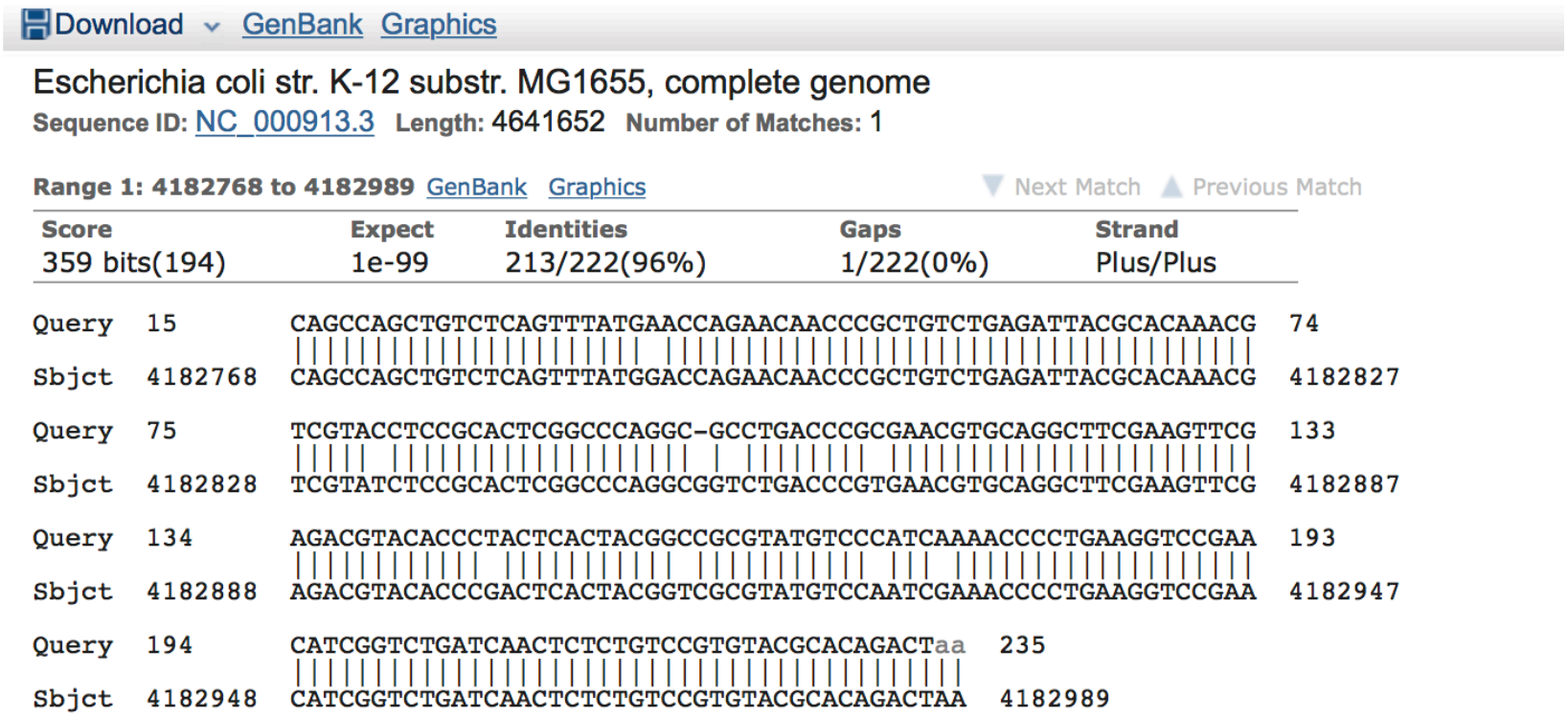
2016/3 AF I

Sequence: 29414192

Samples:	16298	Quality: 0 - 9	
Bases:	307	10 - 19	
Average spacing:	54.0	20 - 29	
Average quality >=	10: 179, 20: 58, 30: 20	>= 30	

22.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182989 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
359 bits(194)	1e-99	213/222(96%)	1/222(0%)	Plus/Plus

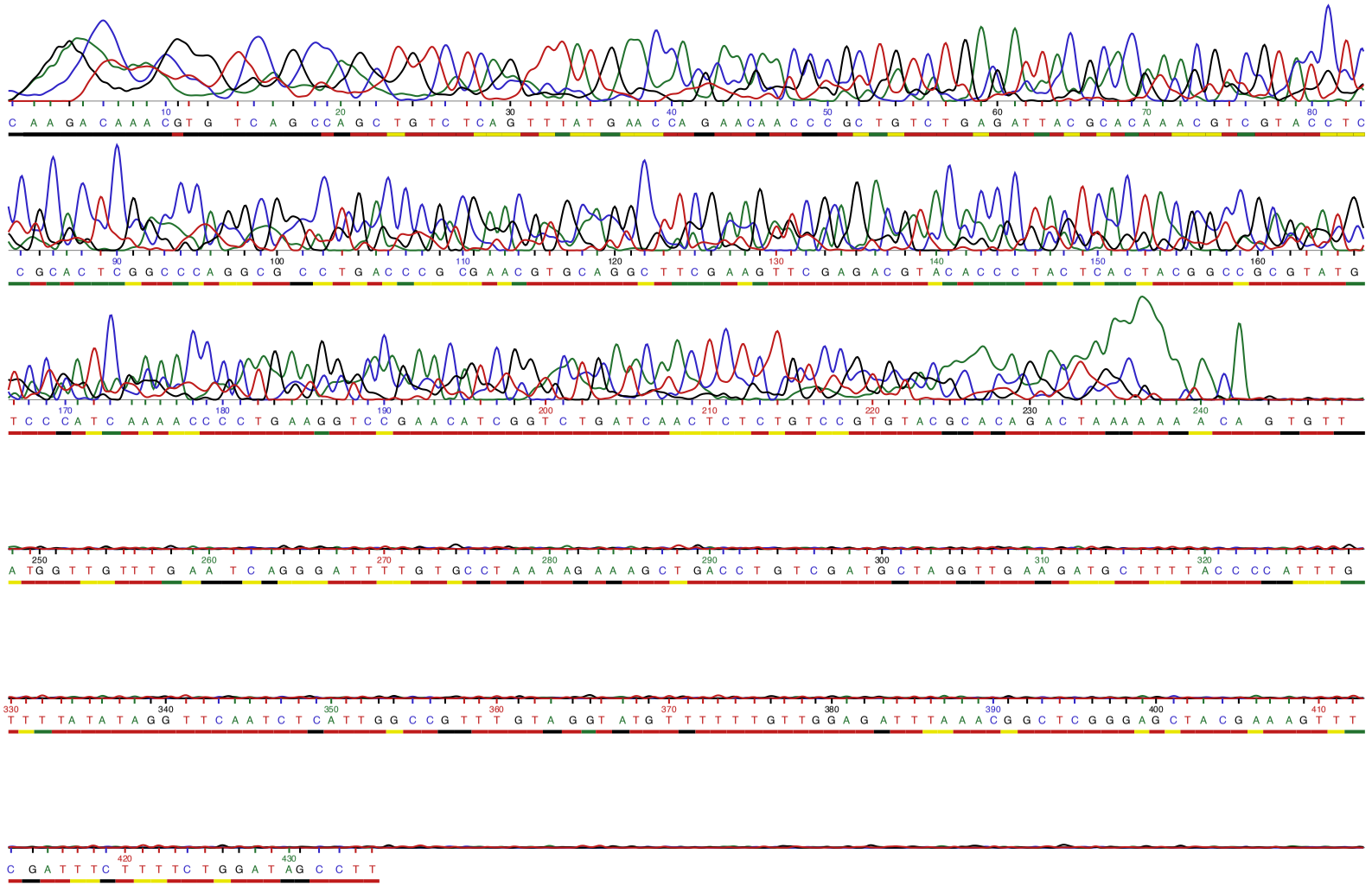
```
Query  15       CAGCCAGCTGTCTCAGTTTATGAACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  74
                ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       TCGTACCTCCGCACTCGGCCCAGGC-GCCTGACCCGCGAACGTGCAGGCTTCGAAGTTCG  133
                ||||| |||||||||||||||||| | ||||||||| ||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCTACTCACTACGGCCGCGTATGTCCCATCAAAACCCCTGAAGGTCCGAA  193
                |||||||||||| |||||||||| |||||||||| ||| ||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTaa  235
                ||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2016/3 AF II

Sequence: 29414277

Samples:	16302	Quality: 0 - 9	
Bases:	436	10 - 19	
Average spacing:	38.0	20 - 29	
Average quality >=	10: 257, 20: 88, 30: 34	>= 30	

Page: 1 / 3
22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182977 GenBank Graphics Next Match Previous Match

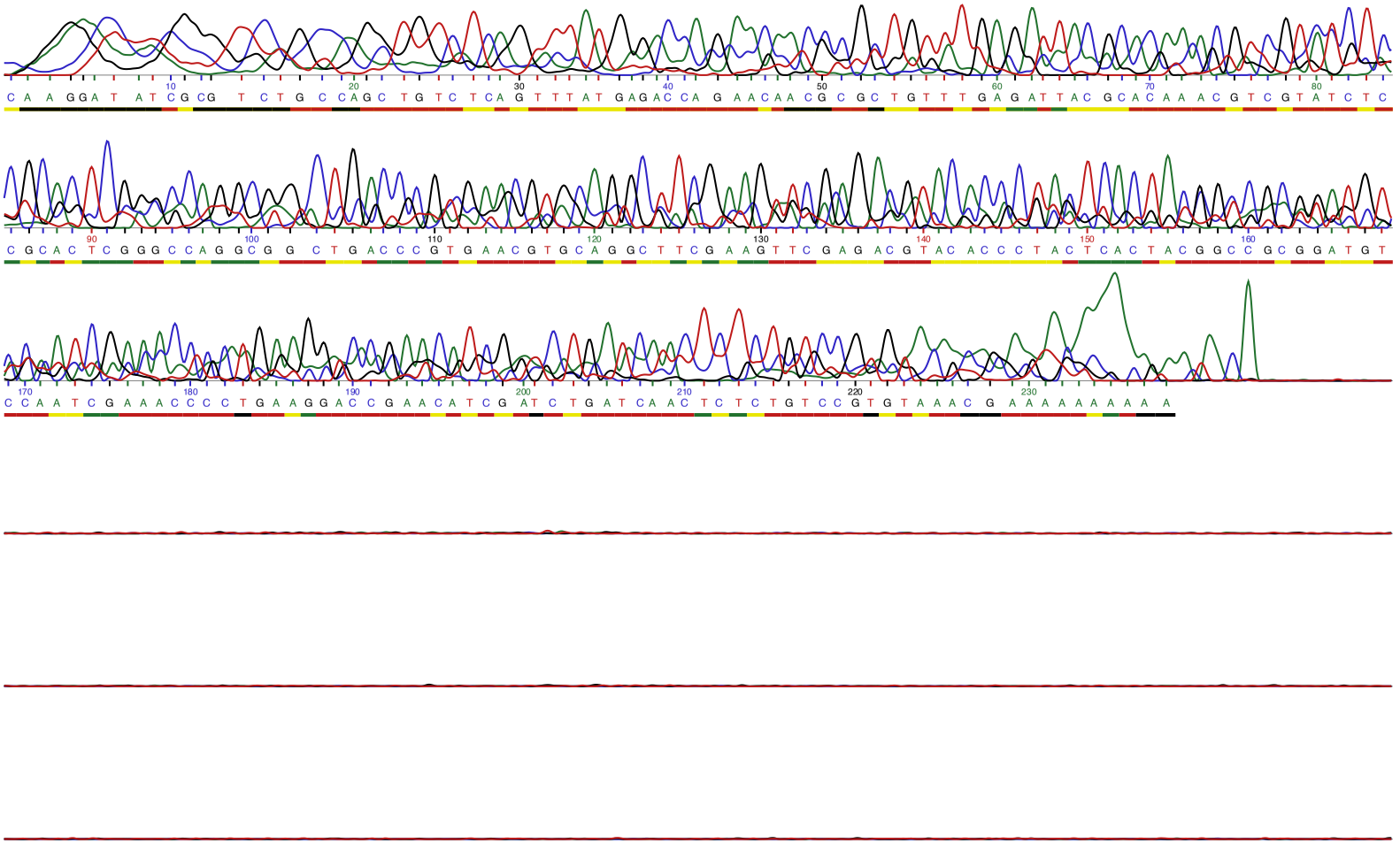
Score	Expect	Identities	Gaps	Strand
329 bits(178)	5e-91	199/209(95%)	2/209(0%)	Plus/Plus

```
Query  17       GCCAGCTGTCTCAGTTTATGAGACCAGAACAACGCGCTGTTTGAGATTACGCACAAACGT  76
                ||||||||||||||||||||| |||||||||||| |||||| ||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  4182828

Query  77       CGTATCTCCGCACTCGGGCCAGGCGG-CTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  135
                |||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||
Sbjct  4182829  CGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  4182888

Query  136      GACGTACACCCTACTCACTACGGCCGCGGATGTCCAATCGAAACCCCTGAAGGACCGAAC  195
                |||||||||||| |||||||||||| |||| |||||||||||||||||||||| ||||||
Sbjct  4182889  GACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAAC  4182948

Query  196      ATCGATCTGATCAACTCTCTGTCCGTGTA  224
                |||| ||||||||||||||||||||||||
Sbjct  4182949  ATCGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

2016/3 AF III

Sequence: 29414352

Samples:	16299	Quality:	0 - 9
Bases:	239		10 - 19
Average spacing:	69.0		20 - 29
Average quality >=	10: 118, 20: 64, 30: 30		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182988 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
370 bits(200)	3e-103	214/221(97%)	0/221(0%)	Plus/Plus

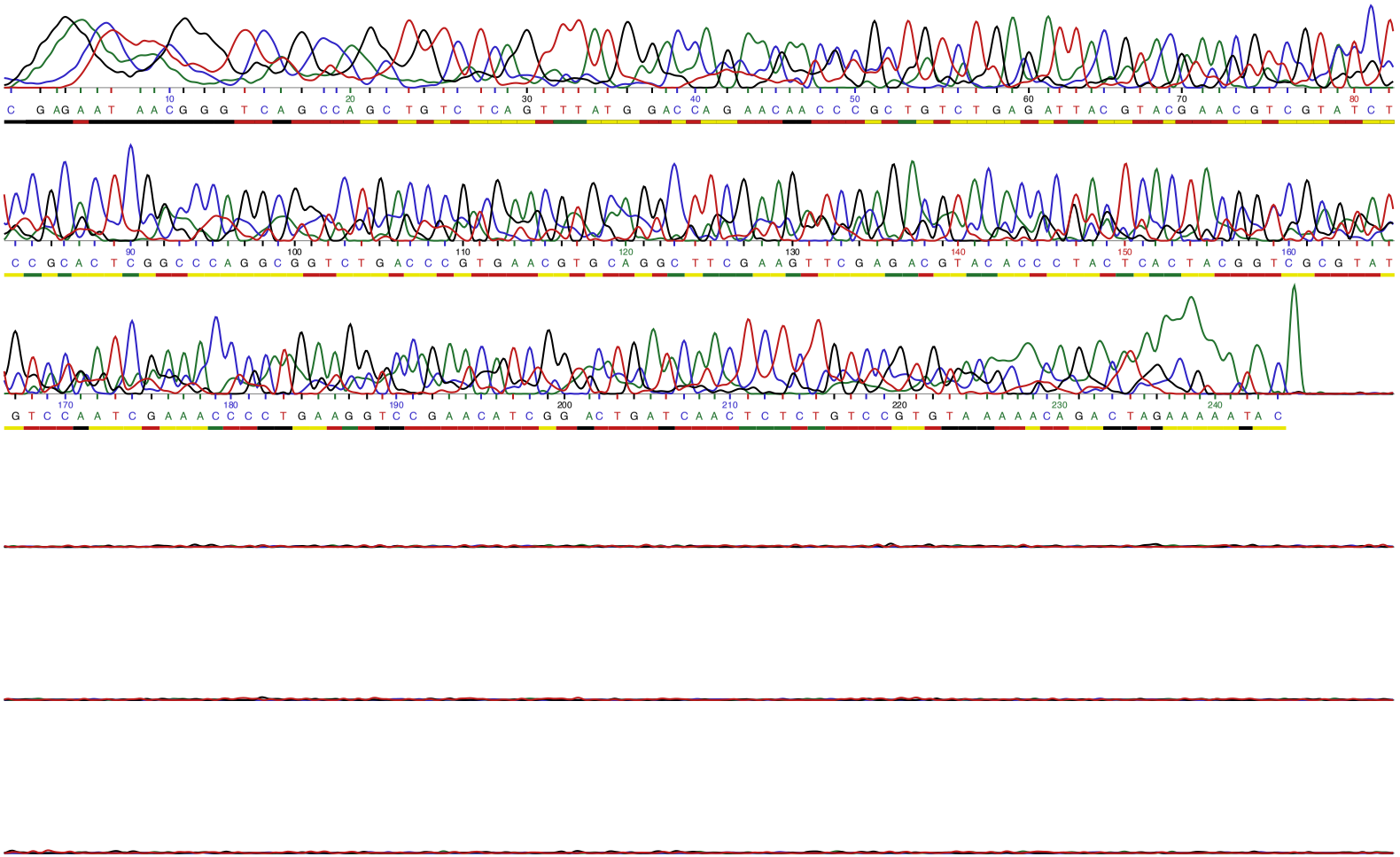
```
Query  15       CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGTACGAACG  74
                |||||||||||||||||||||||||||||||||||||||||||||||||||  ||  ||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  135      AGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  194
                |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  195      CATCGGACTGATCAACTCTCTGTCCGTGTAAAAACAGACTA  235
                ||||||  ||||||||||||||||||||||    ||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTA  4182988
```

2016/3 AF IV

Sequence: 29414468

Samples:	16303	Quality: 0 - 9	
Bases:	245	10 - 19	
Average spacing:	67.0	20 - 29	
Average quality >=	10: 95, 20: 95, 30: 26	>= 30	

Page: 1 / 3
22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182973 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
270 bits(146)	6e-73	187/207(90%)	2/207(0%)	Plus/Plus

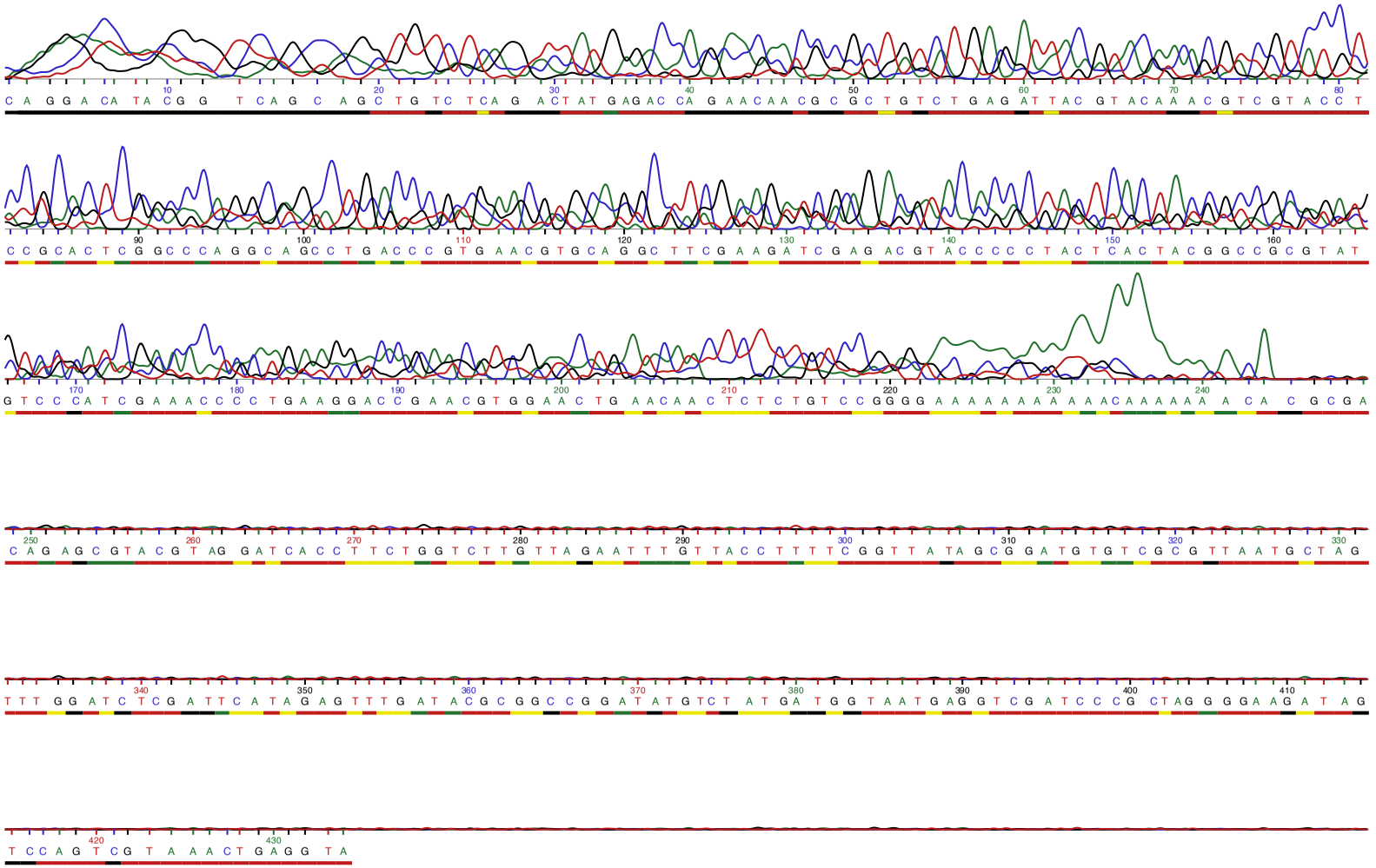
```
Query  14       CAG-CAGCTGTCTCAGACTATGAGACCAGAACAACGCGCTGTCTGAGATTACGTACAAAC  72
               ||| |||||||||||||  ||||| |||||||||||| |||||||||||||||||| ||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  73       GTCGTACCTCCGCACTCGGCCCAGGCAGCCTGACCCGTGAACGTGCAGGCTTCGAAGATC  132
               |||||| |||||||||||||||||| | |||||||||||||||||||||||||||||| ||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  133      GAGACGTACCCCCTACTCACTACGGCCGCGTATGTCCCATCGAAACCCCTGAAGGACCGA  192
               ||||||||| ||| |||||||||||| |||||||||||| |||||||||||||||||| ||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  193      ACGTGGAACTGAACAACTCTCTGTCCG  219
               || | |  |||| ||||||||||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCG  4182973
```

2016/3 AF V

Sequence: 29414482

Samples:	16302	Quality: 0 - 9	
Bases:	435	10 - 19	
Average spacing:	38.0	20 - 29	
Average quality >=	10: 248, 20: 89, 30: 41	>= 30	

22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182966 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
272 bits(147)	1e-73	182/199(91%)	2/199(1%)	Plus/Plus

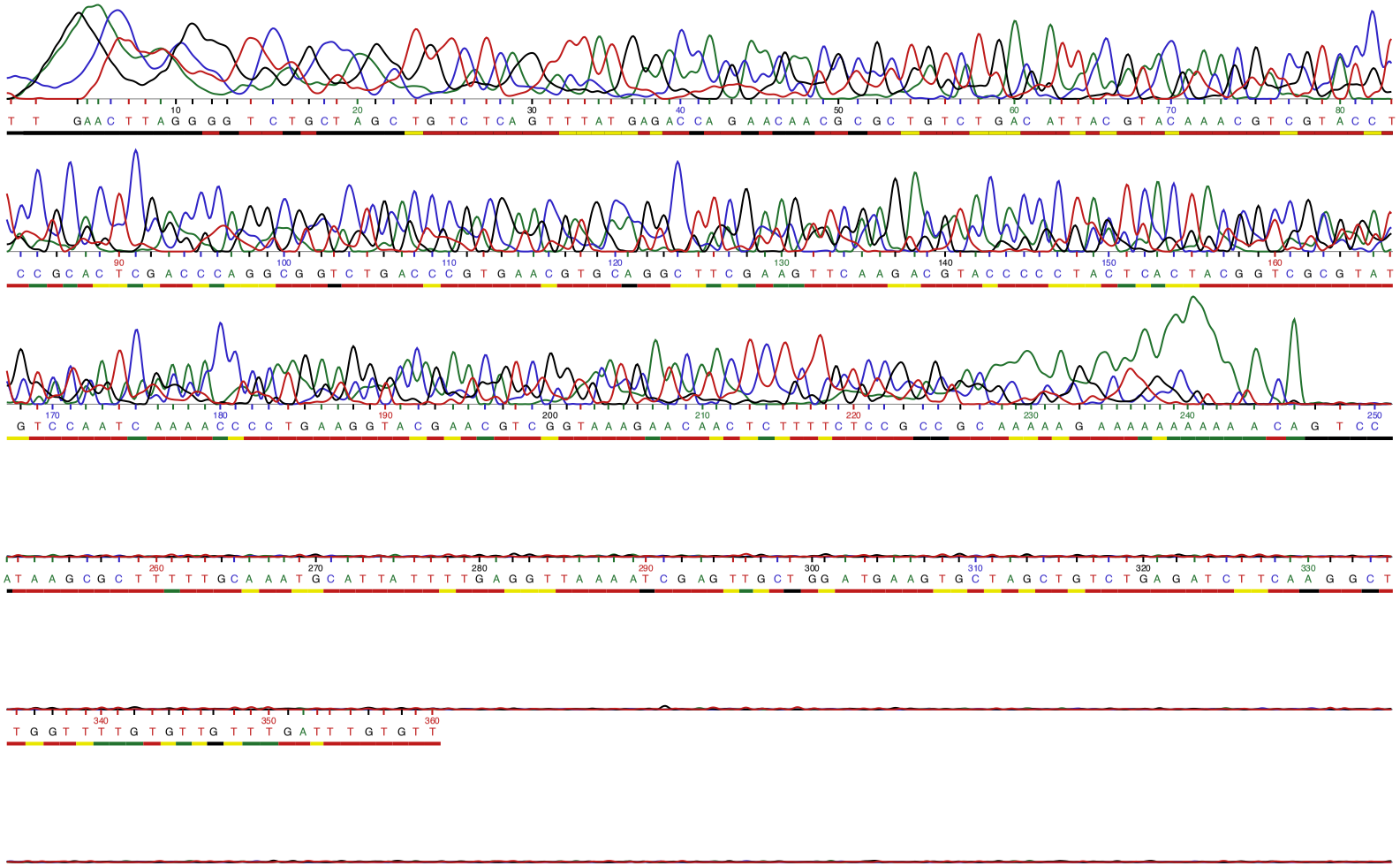
```
Query  17       GCTAGCTGTCTCAGTTTATGAGACCAGAACAACGCGCTGTCTGACATTACGTACAAACGT  76
                || |||||||||||||||| |||||||||||| |||||||||| |||||| ||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGT  4182828

Query  77       CGTACCTCCGCACTCGACCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCAA  136
                ||||  |||||||||||| ||||||||||||||||||||||||||||||||||||||| |
Sbjct  4182829  CGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGA  4182888

Query  137      GACGTACCCCCTACTCACTACGGTCGCGTATGTCCAATCAAAACCCCTGAAGGTACGAAC  196
                |||||||  |||  |||||||||||||||||||||||||| ||||||||||||||  |||||
Sbjct  4182889  GACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAAC  4182948

Query  197      GTCGGTAAAGAACAACTCT  215
                 |||||   || |||||||
Sbjct  4182949  ATCGGTCT-GATCAACTCT  4182966
```

2016/3 BK III

Sequence: 29414604

Samples:	16302	Quality: 0 - 9	
Bases:	361	10 - 19	
Average spacing:	46.0	20 - 29	
Average quality >=	10: 219, 20: 73, 30: 31	>= 30	

22.05.2019



2016/3 BK IV

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182760 to 4182933 GenBank Graphics Next Match Previous Match

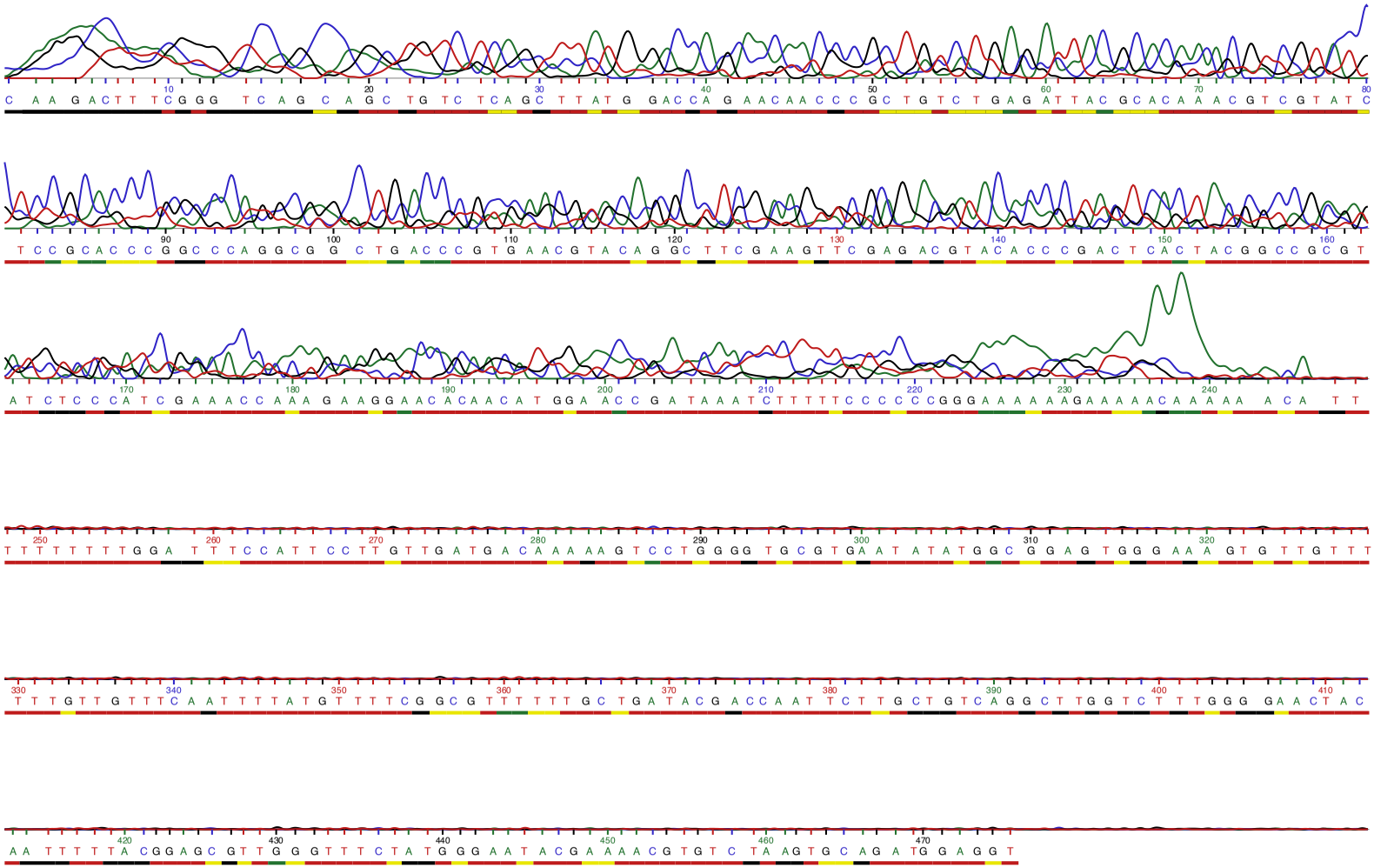
Score	Expect	Identities	Gaps	Strand
263 bits(142)	1e-70	164/174(94%)	3/174(1%)	Plus/Plus

```
Query  8        TTCGGGT-CAG-CAGCTGTCTCAGCTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  65
               ||||| | ||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  4182760  TTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  4182819

Query  66       CACAAACGTCGTATCTCCGCACCCGGCCCAGGCGG-CTGACCCGTGAACGTACAGGCTTC  124
               |||||||||||||||||||||| |||||||||||| |||||||||||||| ||||||||
Sbjct  4182820  CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  4182879

Query  125      GAAGTTCGAGACGTACACCCGACTCACTACGGCCGCGTATCTCCCATCGAAACC  178
               |||||||||||||||||||||||||||||||| ||||||| ||| ||||||||||
Sbjct  4182880  GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACC  4182933
```

2016/3 BK IV

Sequence: 29414642

Samples:	16303	Quality:	0 - 9
Bases:	476		10 - 19
Average spacing:	35.0		20 - 29
Average quality >=	10: 302, 20: 82, 30: 22		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182973 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
329 bits(178)	1e-90	197/206(96%)	2/206(0%)	Plus/Plus

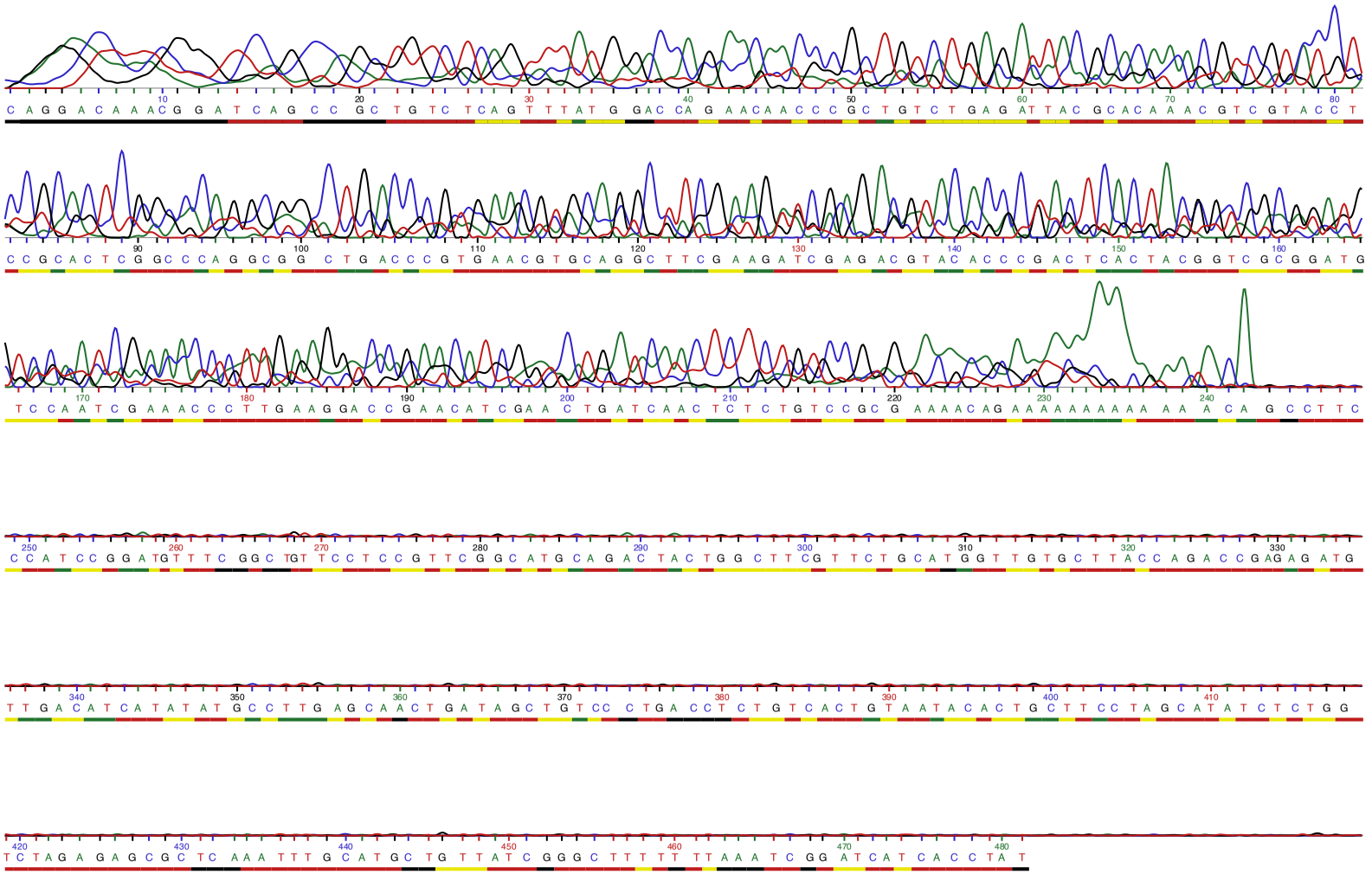
```
Query  15       CAGCC-GCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  73
                |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  74       TCGTACCTCCGCACTCGGCCCAGGCGG-CTGACCCGTGAACGTGCAGGCTTCGAAGATCG  132
                |||||  |||||||||||||||||||| |||||||||||||||||||||||||||  |||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  133      AGACGTACACCCGACTCACTACGGTCGCGGATGTCCAATCGAAACCCTTGAAGGACCGAA  192
                |||||||||||||||||||||||||||| |||||||||||||||| |||||| |||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  193      CATCGAACTGATCAACTCTCTGTCCG  218
                |||||  |||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCG  4182973
```

2016/3 BK V

Sequence: 29414666

Samples:	16302	Quality: 0 - 9	
Bases:	482	10 - 19	
Average spacing:	34.0	20 - 29	
Average quality >=	10: 235, 20: 143, 30: 60	>= 30	

Page: 1 / 3
22.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182985 GenBank Graphics Next Match Previous Match

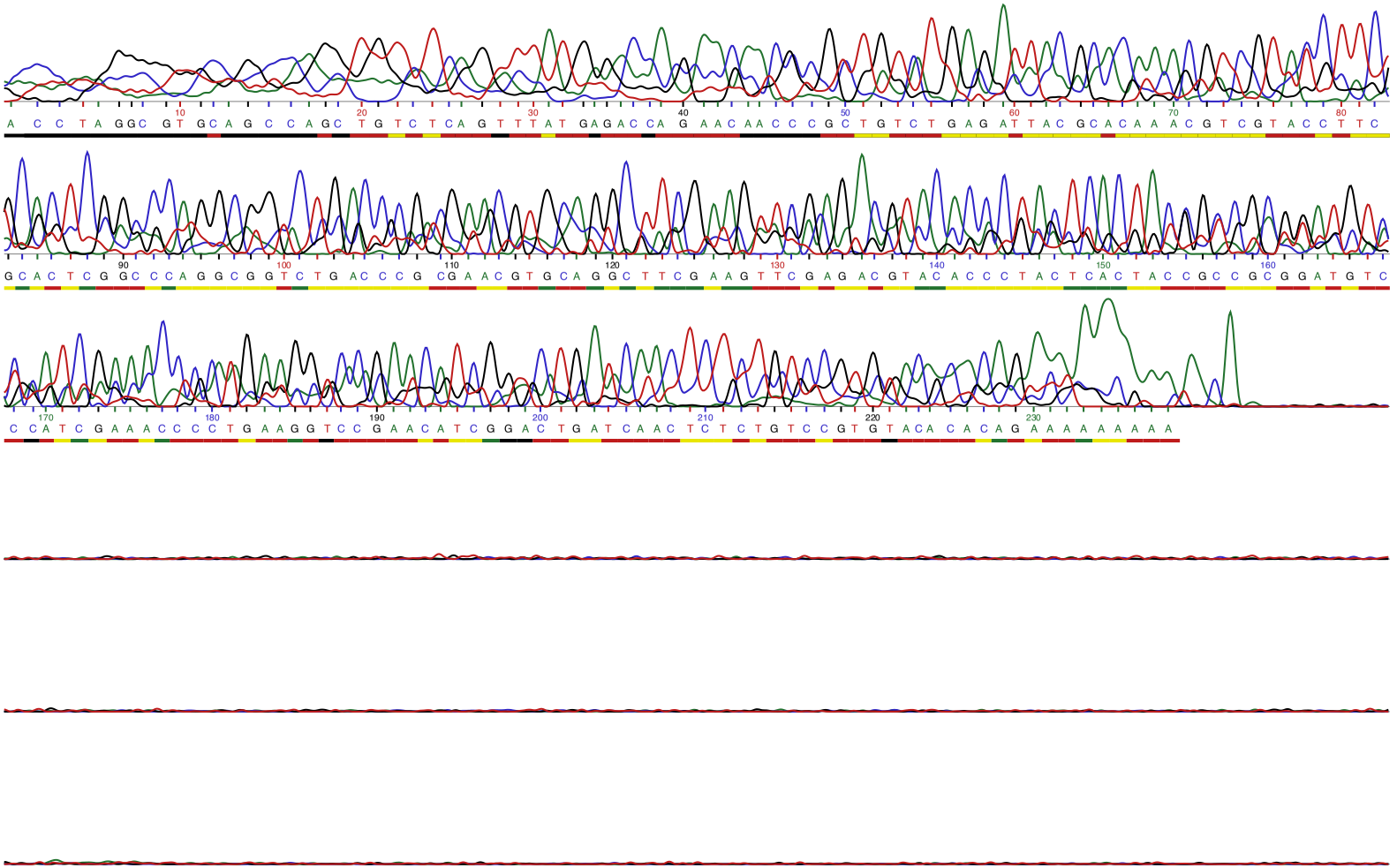
Score	Expect	Identities	Gaps	Strand
342 bits(185)	7e-95	208/219(95%)	1/219(0%)	Plus/Plus

```
Query  12       CAGCCAGCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  71
                |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  72       GTCGTACCTTCGCACTCGGCCCAGGCGGTCTGACCCGCGAACGTGCAGGCTTCGAAGTTC  131
                ||||||  || ||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  132      GAGACGTACACCCTACTCACTACCGCCGCGGATGTCCCATCGAAACCCCTGAAGGTCCGA  191
                |||||||||||||  |||||||||  | ||||  ||||||  ||||||||||||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  192      ACATCGGACTGATCAACTCTCTGTCCGTGTACACACAGa  230
                |||||||  |||||||||||||||||||||||||  ||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2016/3 BF I

Sequence: 29416196

Samples:	16301	Quality:	0 - 9
Bases:	239		10 - 19
Average spacing:	69.0		20 - 29
Average quality >=	10: 99, 20: 85, 30: 24		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182772 to 4182980 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
257 bits(139)	7e-69	186/209(89%)	1/209(0%)	Plus/Plus

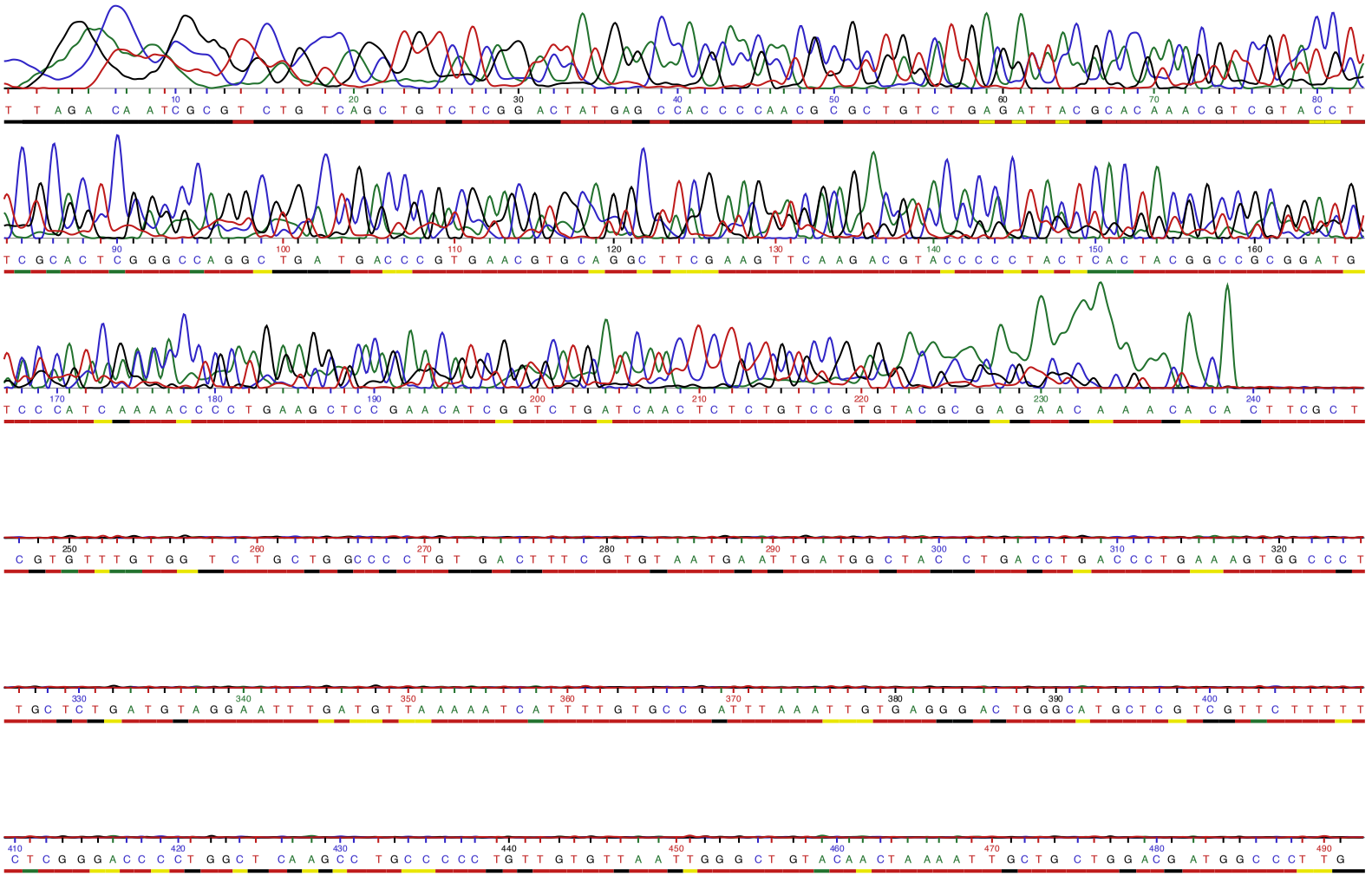
```
Query  19       CAGCTGTCTCGGACTATGAGCCACCCCAACGCGCTGTCTGAGATTACGCACAAACGTCGT  78
                ||||||||||| |  ||||  |||   ||||| ||||||||||||||||||||||||||
Sbjct  4182772  CAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGT  4182831

Query  79       ACCTTCGCACTCGGGCCAGGCTG-ATGACCCGTGAACGTGCAGGCTTCGAAGTTCAAGAC  137
                | || |||||||||  ||||||  |  ||||||||||||||||||||||||||||| ||||
Sbjct  4182832  ATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGAC  4182891

Query  138      GTACCCCCTACTCACTACGGCCGCGGATGTCCCATCAAAACCCCTGAAGCTCCGAACATC  197
                ||||  |||  ||||||||||||  ||||  ||||||  |||  ||||||||||||  ||||||||||
Sbjct  4182892  GTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATC  4182951

Query  198      GGTCTGATCAACTCTCTGTCCGTGTACGC  226
                |||||||||||||||||||||||||||||
Sbjct  4182952  GGTCTGATCAACTCTCTGTCCGTGTACGC  4182980
```

2016/3 BF III

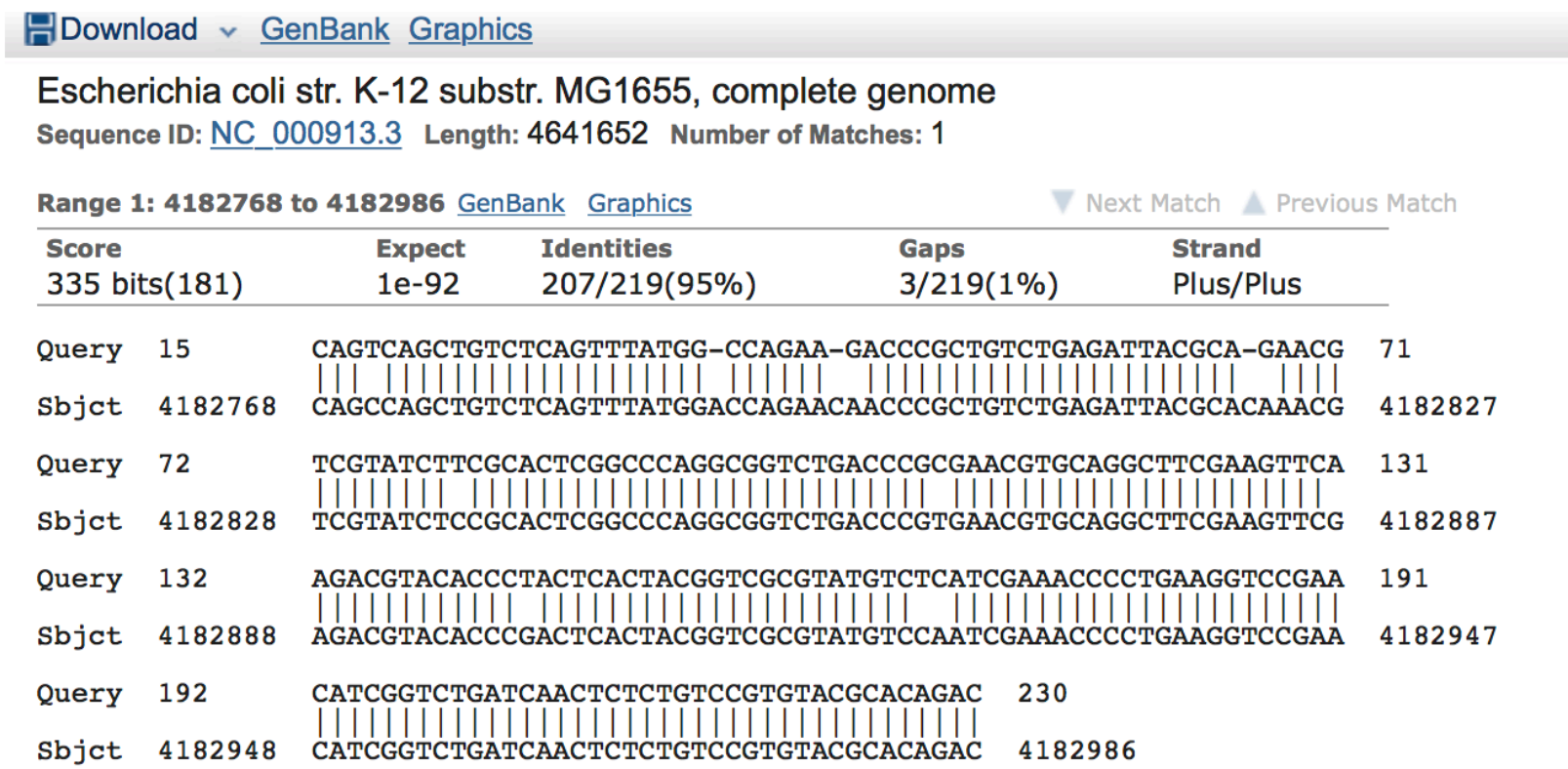
Sequence: 29414987

Samples:	16304	Quality: 0 - 9	
Bases:	656	10 - 19	
Average spacing:	25.0	20 - 29	
Average quality >=	10: 448, 20: 63, 30: 14	>= 30	

22.05.2019



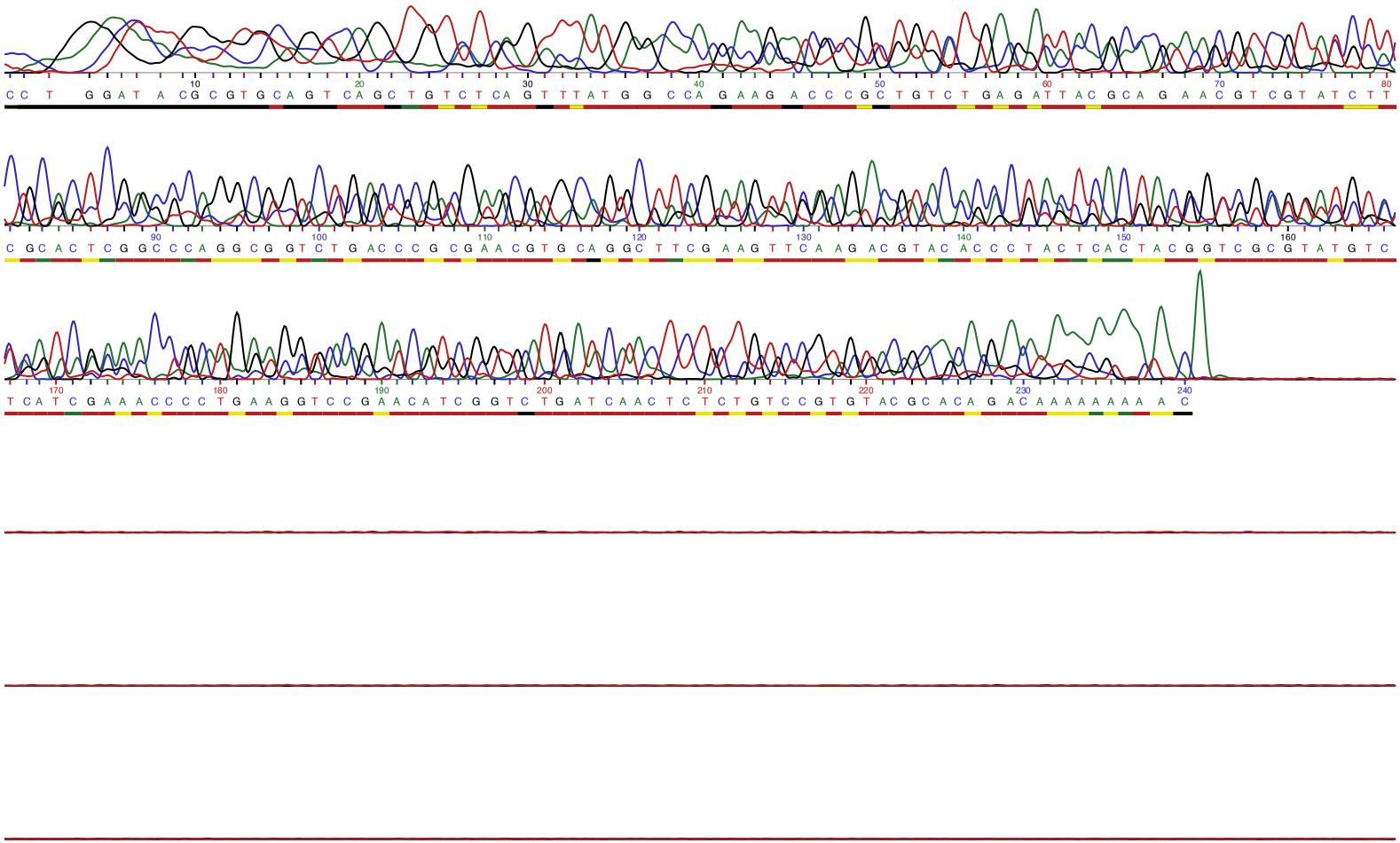
2016/4 AK IV


Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182768 to 4182986 GenBank Graphics
Next Match Previous Match
Score 335 bits(181)
Expect 1e-92
Identities 207/219(95%)
Gaps 3/219(1%)
Strand Plus/Plus
Query 15 CAGTCAGCTGTCTCAGTTTATGG-CCAGAA-GACCCGCTGTCTGAGATTACGCA-GAACG 71
Sbjct 4182768 CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG 4182827
Query 72 TCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGCGAACGTGCAGGCTTCGAAGTTCA 131
Sbjct 4182828 TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG 4182887
Query 132 AGACGTACACCCTACTCACTACGGTCGCGTATGTCTCATCGAAACCCCTGAAGGTCCGAA 191
Sbjct 4182888 AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA 4182947
Query 192 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC 230
Sbjct 4182948 CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC 4182986

2016/4 AK IV

Sequence: 29412839

Samples:	15428	Quality:	0 - 9
Bases:	241		10 - 19
Average spacing:	65.0		20 - 29
Average quality >=	10: 150, 20: 53, 30: 13		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
231 bits(125)	2e-61	188/219(86%)	2/219(0%)	Plus/Plus

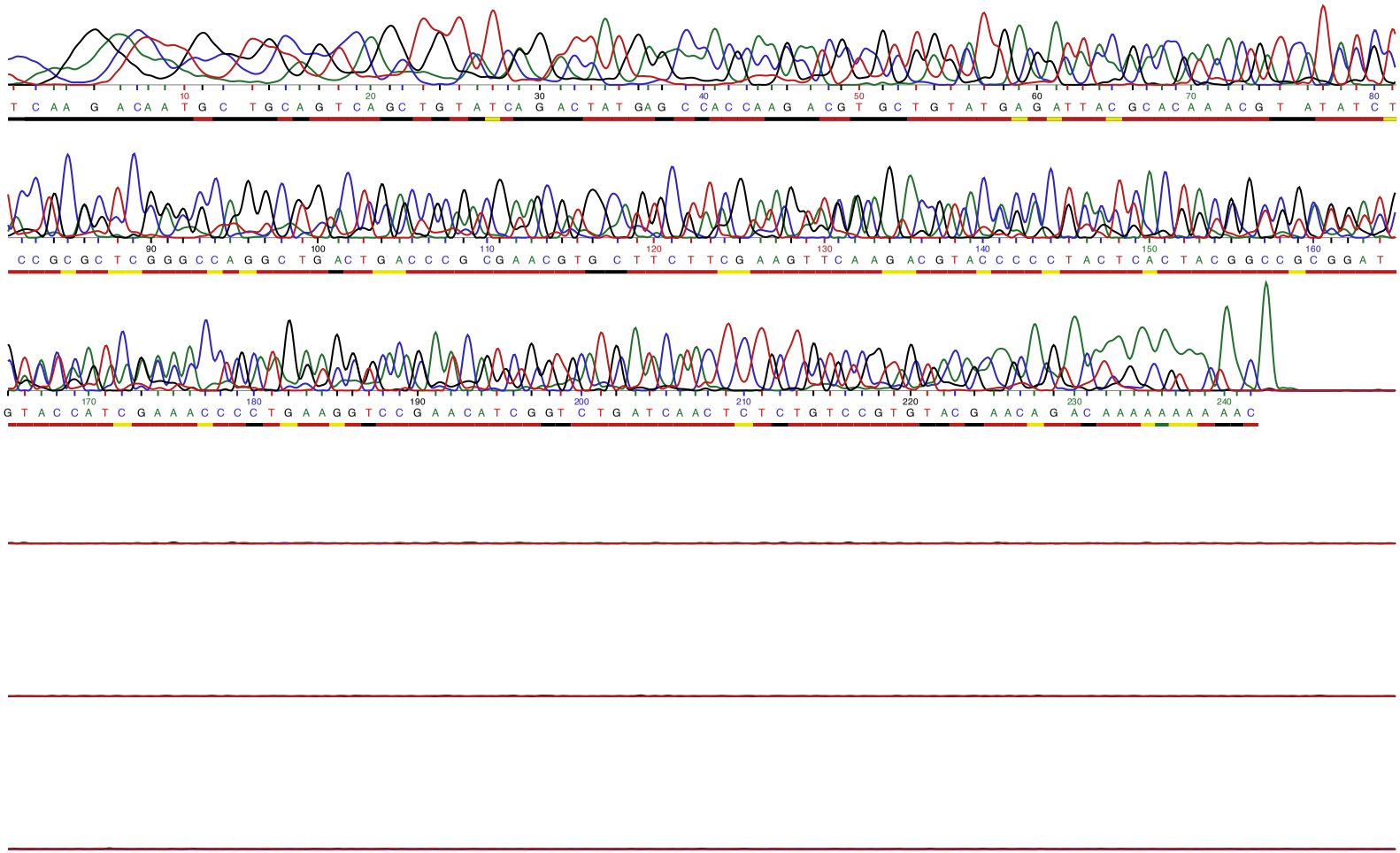
```
Query  15       CAGTCAGCTGTATCAGACTATGAGCCACCAAGACGTGCTGTATGAGATTACGCACAAACG  74
                |||  |||||||  ||||  ||||  |||  |  ||  |||||  ||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  75       T-ATATCTCCGCGCTCGGGCCAGGCTGACTGACCCGCGAACGTGC-TTCTTCGAAGTTCA  132
                |  |||||||||  |||||  |||||||  |  ||||||||  ||||||||   ||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  133      AGACGTACCCCCTACTCACTACGGCCGCGGATGTACCATCGAAACCCCTGAAGGTCCGAA  192
                ||||||||  |||  ||||||||||||  ||||  ||||  |  ||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  193      CATCGGTCTGATCAACTCTCTGTCCGTGTACGAACAGAC  231
                ||||||||||||||||||||||||||||||||||||||  |||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

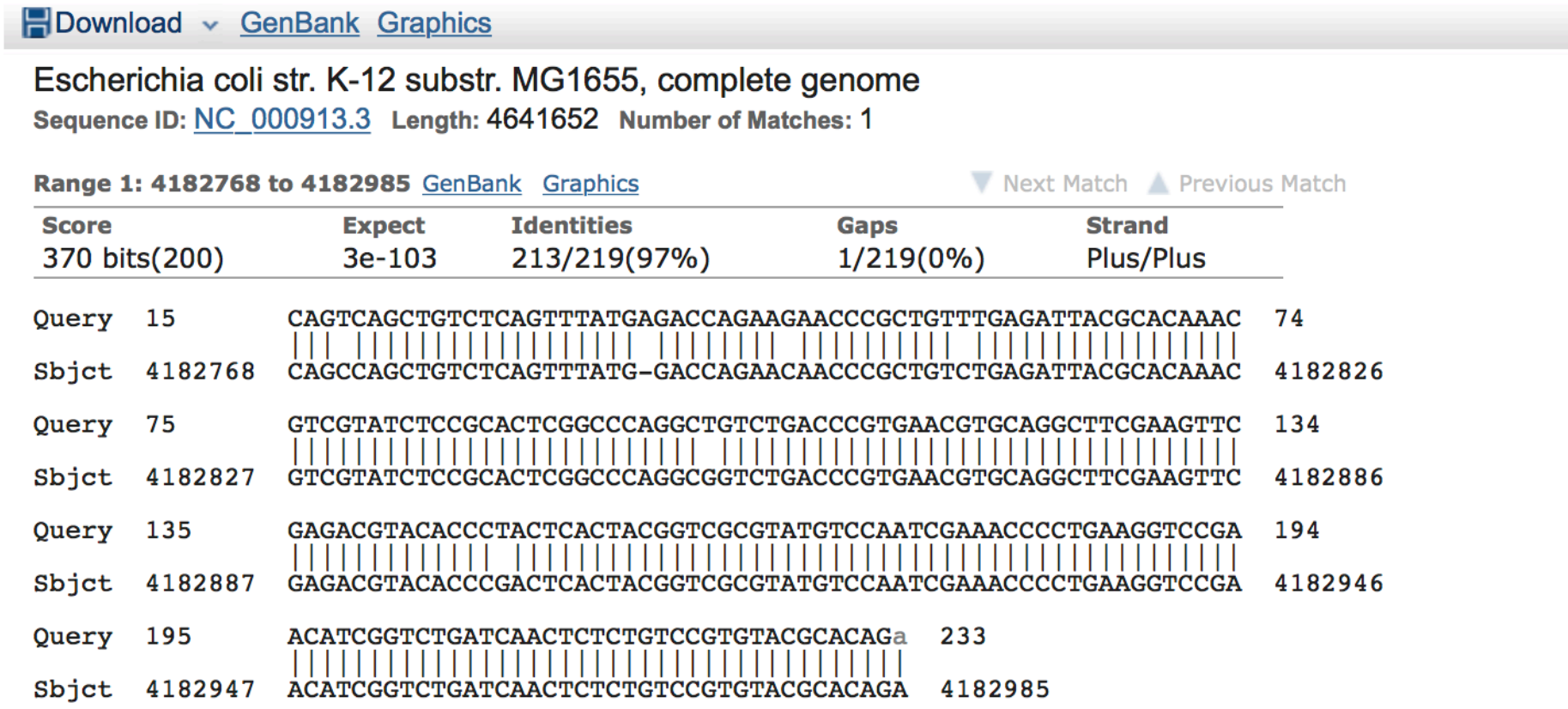
2016/4 AF I

Sequence: 29412877

Samples:	15219	Quality:	0 - 9
Bases:	243		10 - 19
Average spacing:	63.0		20 - 29
Average quality >=	10: 166, 20: 30, 30: 1		>= 30

Page: 1 / 3
16.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182985 GenBank Graphics Next Match Previous Match

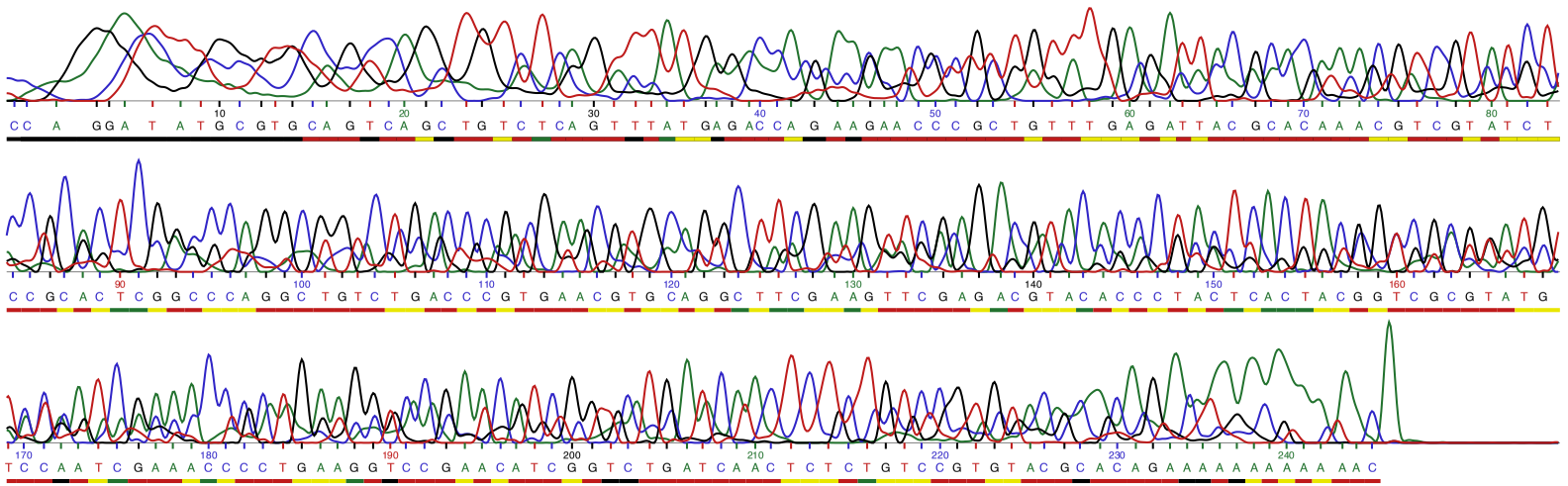
Score	Expect	Identities	Gaps	Strand
370 bits(200)	3e-103	213/219(97%)	1/219(0%)	Plus/Plus

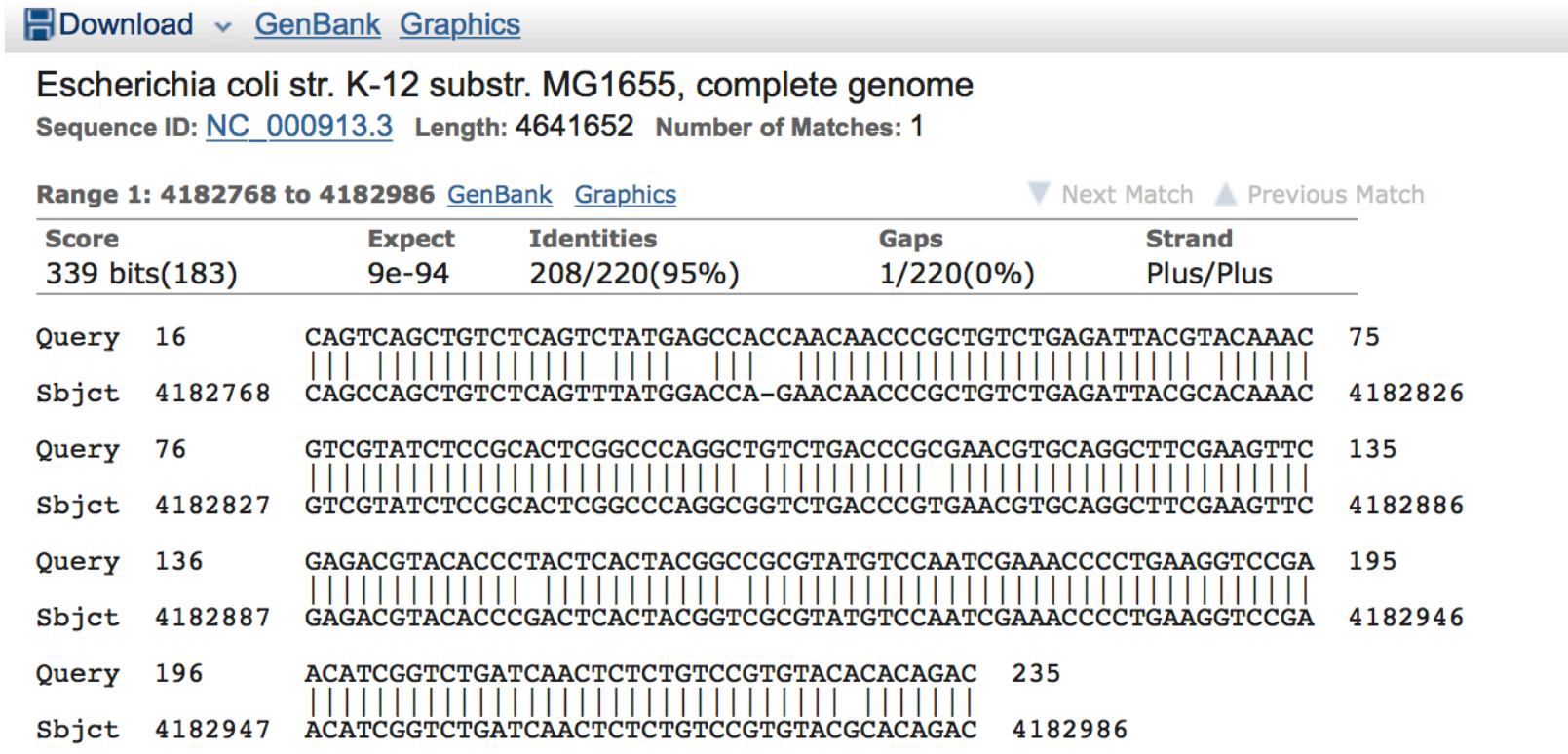
```
Query  15       CAGTCAGCTGTCTCAGTTTATGAGACCAGAAGAACCCGCTGTTTGAGATTACGCACAAAC  74
                |||  |||||||||||||||||||  ||||||||  ||||||||||  ||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  75       GTCGTATCTCCGCACTCGGCCCAGGCTGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  134
                ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  135      GAGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  194
                |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  195      ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGa  233
                |||||||||||||||||||||||||||||||||||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2016/4 AF IV

Sequence: 29412976

Samples:	15281	Quality:	0 - 9
Bases:	246		10 - 19
Average spacing:	63.0		20 - 29
Average quality >=	10: 130, 20: 70, 30: 18		>= 30





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182986 GenBank Graphics Next Match Previous Match

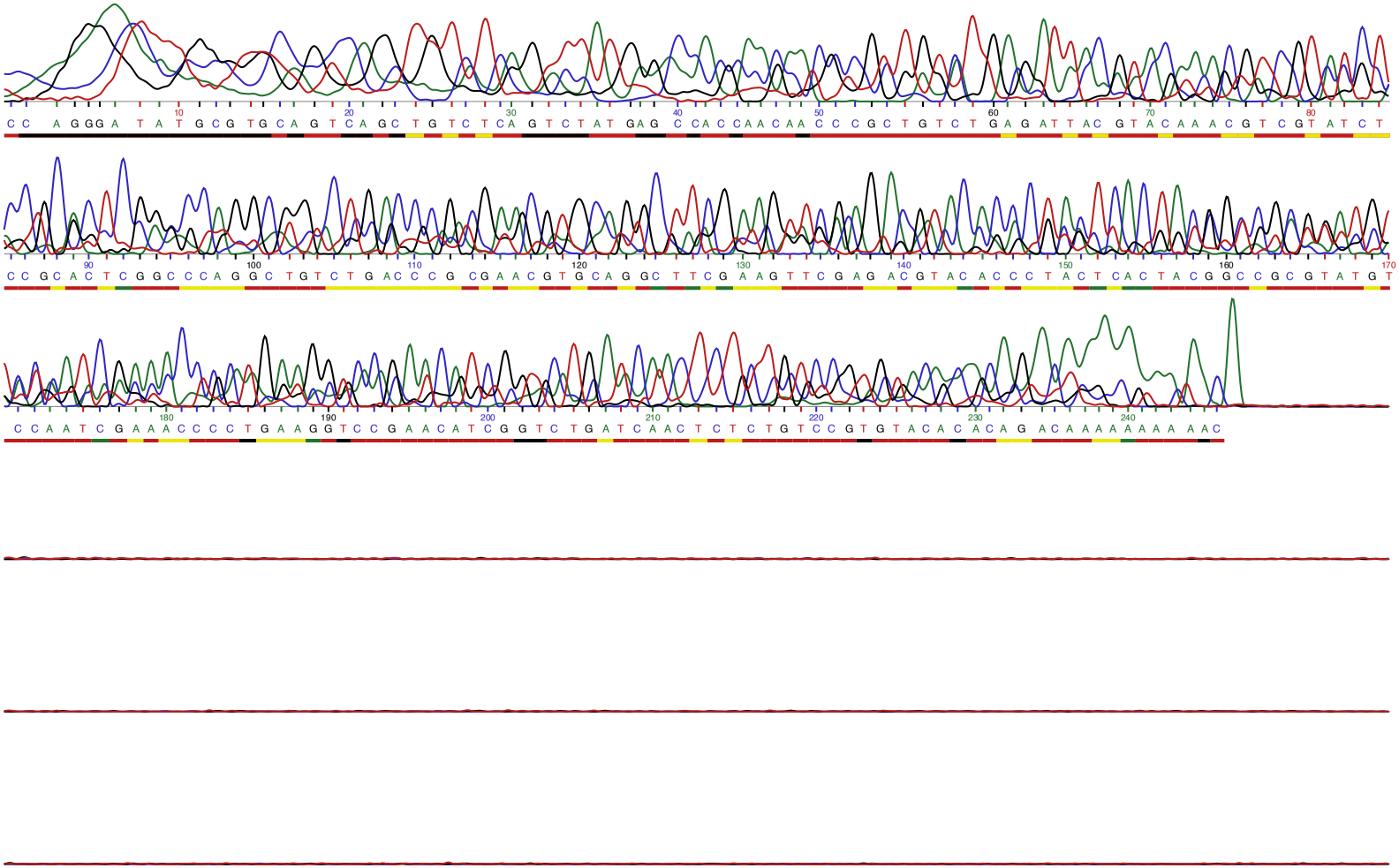
Score	Expect	Identities	Gaps	Strand
339 bits(183)	9e-94	208/220(95%)	1/220(0%)	Plus/Plus

```
Query  16       CAGTCAGCTGTCTCAGTCTATGAGCCACCAACAACCCGCTGTCTGAGATTACGTACAAAC  75
                |||  |||||||||||||| ||||  |||  |||||||||||||||||||||||| ||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCA-GAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  76       GTCGTATCTCCGCACTCGGCCCAGGCTGTCTGACCCGCGAACGTGCAGGCTTCGAAGTTC  135
                |||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  136      GAGACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  195
                |||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  196      ACATCGGTCTGATCAACTCTCTGTCCGTGTACACACAGAC  235
                |||||||||||||||||||||||||||||||| |||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

2016/4 BK II

Sequence: 29413201

Samples:	15301	Quality:	0 - 9
Bases:	247		10 - 19
Average spacing:	62.0		20 - 29
Average quality >=	10: 141, 20: 61, 30: 11		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
337 bits(182)	3e-93	207/219(95%)	1/219(0%)	Plus/Plus

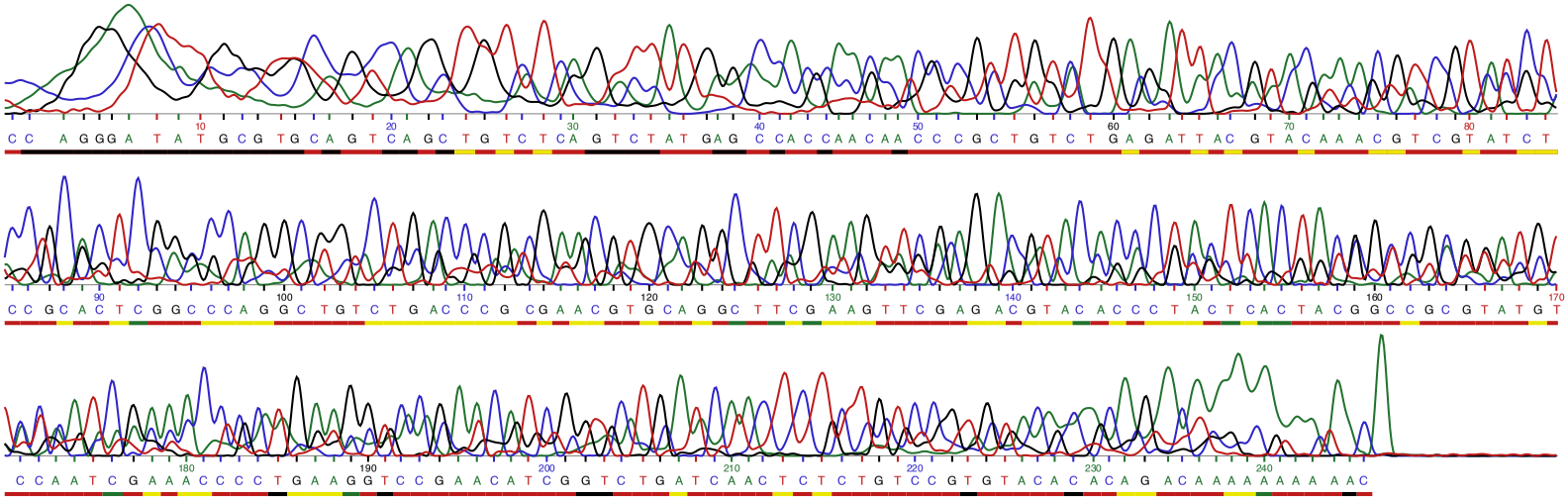
```
Query  15       CAGTCAGGTGTCTCTGTTTATGG-CCACCAAGACCCGCTGTCTGAGATTACGCACAAACG  73
                |||  |||  ||||||  ||||||||  |||   |  ||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  74       TCGTATCTCCGCGCTCGGGCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  133
                ||||||||||||  |||||  ||||||||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  134      AGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCGAA  193
                ||||||||||||  ||||||||||||||||||||||  ||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  194      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  232
                ||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

2016/4 BK III

Sequence: 29413201

Samples:	15301	Quality: 0 - 9	
Bases:	247	10 - 19	
Average spacing:	62.0	20 - 29	
Average quality >=	10: 141, 20: 61, 30: 11	>= 30	

Page: 1 / 3
16.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
342 bits(185)	7e-95	208/219(95%)	1/219(0%)	Plus/Plus

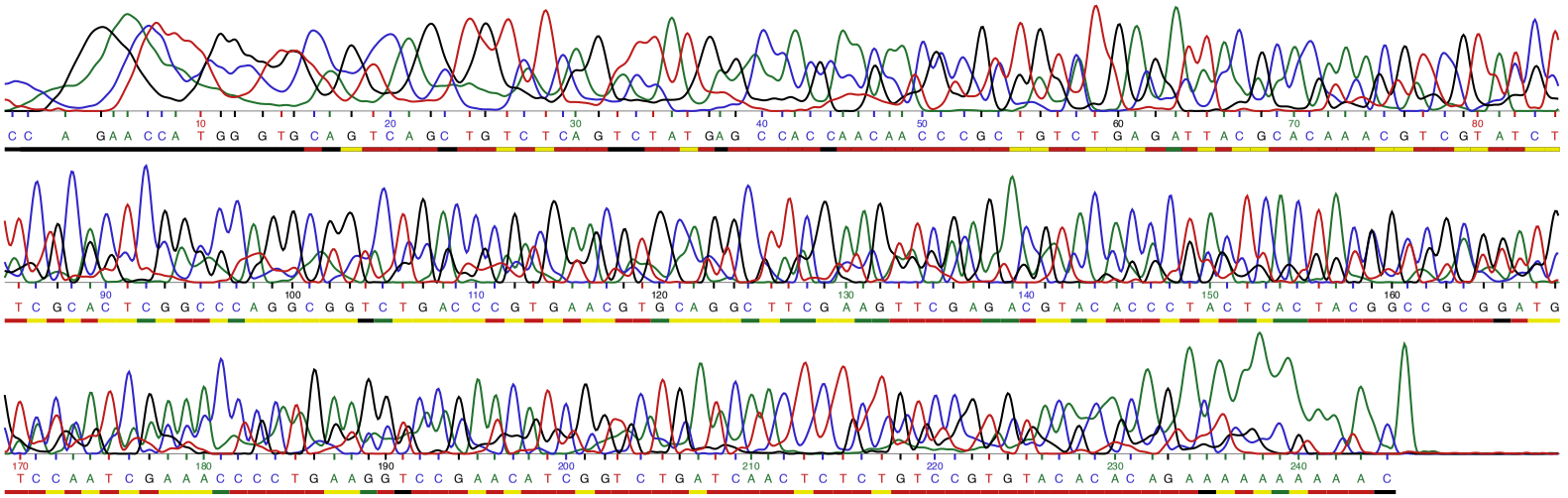
```
Query  16       CAGTCAGCTGTCTCAGTCTATGAGCCACCAACAACCCGCTGTCTGAGATTACGCACAAAC  75
                |||  |||||||||||||  ||||  |||  ||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCA-GAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  76       GTCGTATCTTCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  135
                |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  136      GAGACGTACACCCTACTCACTACGGCCGCGGATGTCCAATCGAAACCCCTGAAGGTCCGA  195
                |||||||||||||  ||||||||||  ||||  ||||||||||||||||||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  196      ACATCGGTCTGATCAACTCTCTGTCCGTGTACACACAGa  234
                ||||||||||||||||||||||||||||||||||||||  ||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2016/4 BK IV

Sequence: 29413409

Samples:	15257	Quality: 0 - 9	
Bases:	245	10 - 19	
Average spacing:	63.0	20 - 29	
Average quality >=	10: 129, 20: 71, 30: 19	>= 30	

Page: 1 / 3
16.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182986 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
359 bits(194)	7e-100	211/219(96%)	2/219(0%)	Plus/Plus

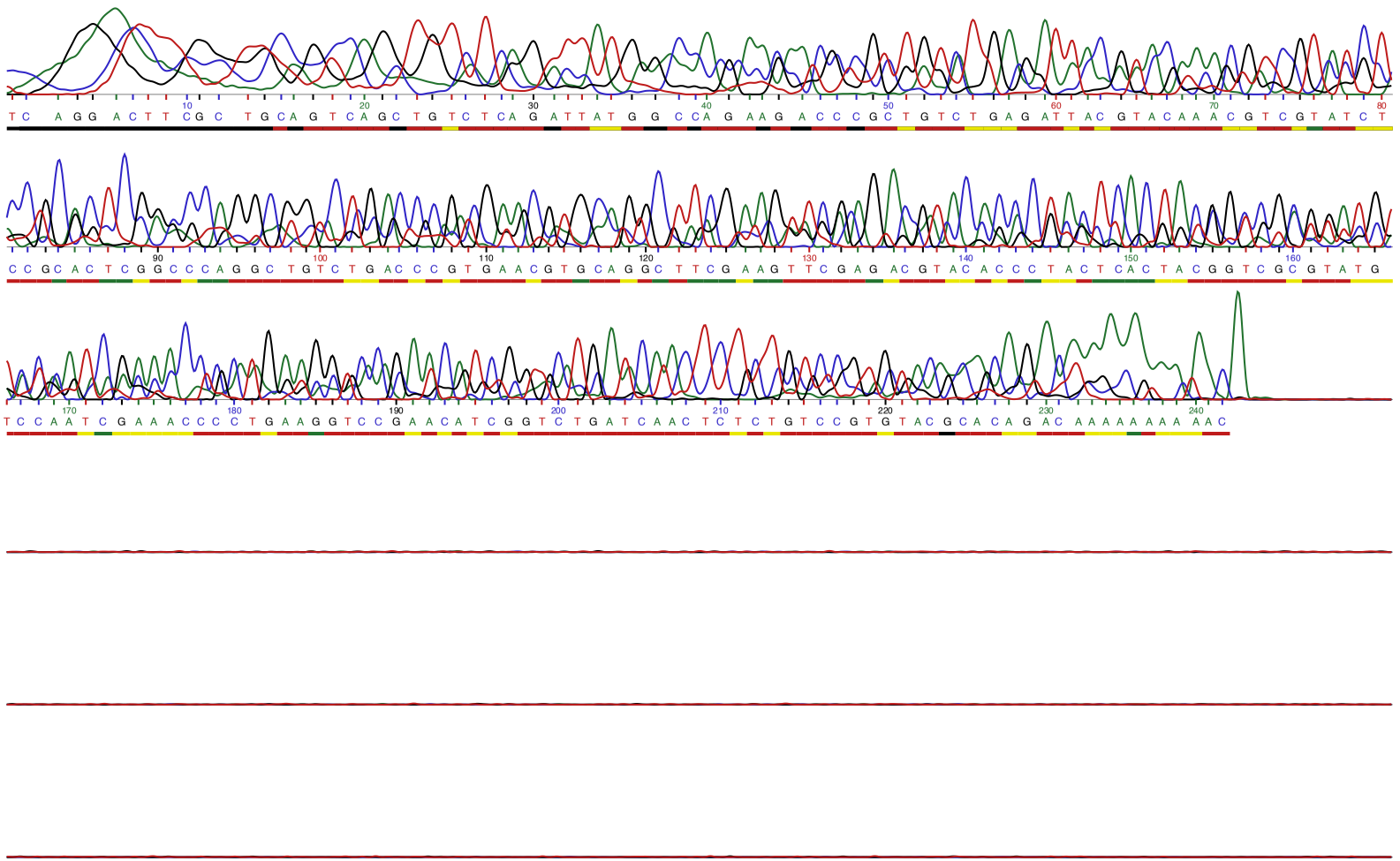
```
Query  15       CAGTCAGCTGTCTCAGATTATGG-CCAGAA-GACCCGCTGTCTGAGATTACGTACAAACG  72
                |||  ||||||||||||  ||||||  ||||||  ||||||||||||||||||||  |||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  73       TCGTATCTCCGCACTCGGCCCAGGCTGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  132
                ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  133      AGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  192
                ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  193      CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  231
                |||||||||||||||||||||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGAC  4182986
```

2016/4 BK V

Sequence: 29413577

Samples:	15313	Quality: 0 - 9	
Bases:	243	10 - 19	
Average spacing:	64.0	20 - 29	
Average quality >=	10: 141, 20: 57, 30: 22	>= 30	

Page: 1 / 3
16.05.2019



Download GenBank Graphics

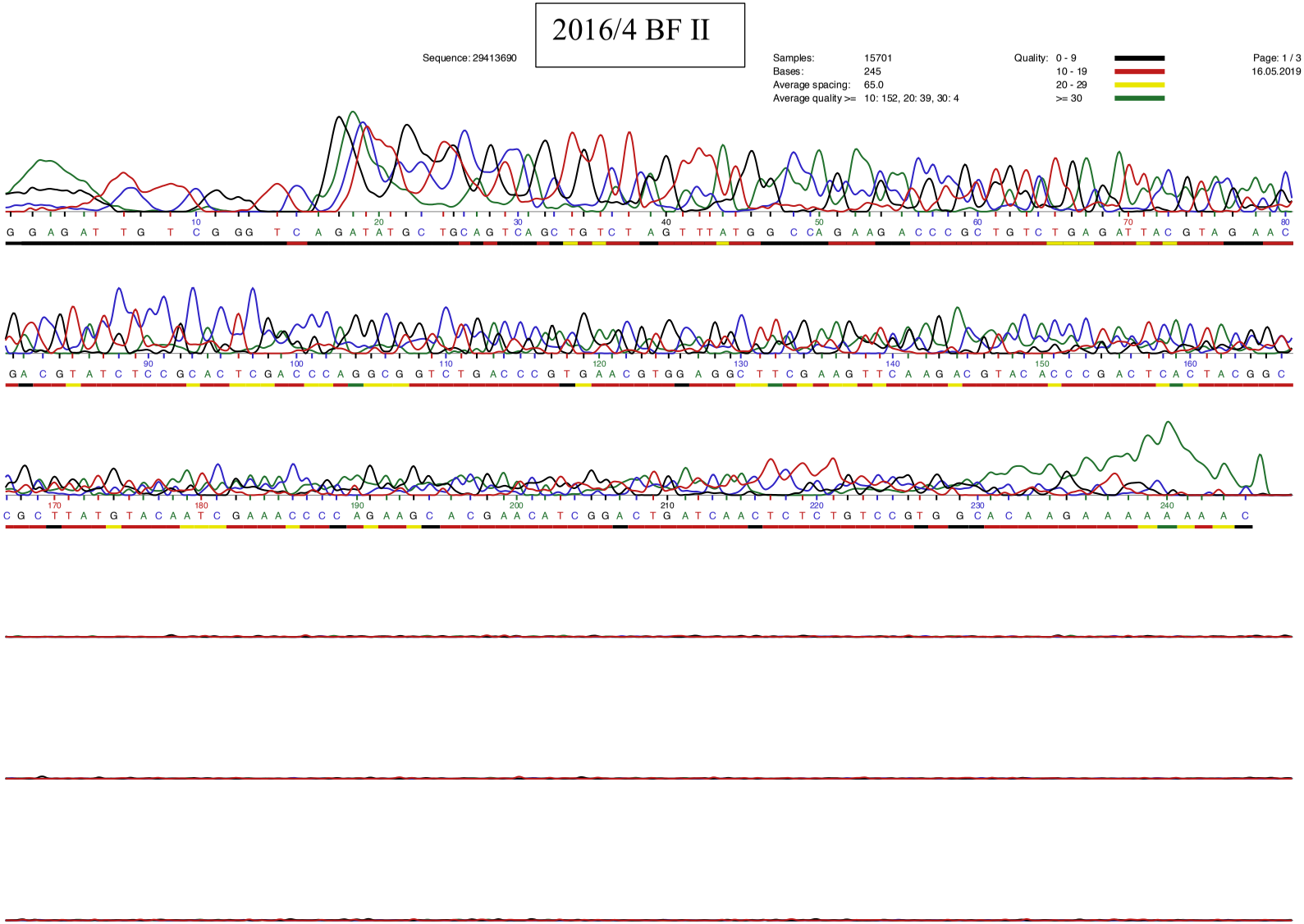
Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182975 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
268 bits(145)	1e-72	188/208(90%)	5/208(2%)	Plus/Plus

```
Query  26       CAGTCAGCTGTCT-AGTTTATGG-CCAGAA-GACCCGCTGTCTGAGATTACGTA-GAACG  81
                |||  |||||||||  |||||||||  ||||||  ||||||||||||||||||||  |  ||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACG  4182827

Query  82       ACGTATCTCCGCACTCGACCCAGGCGGTCTGACCCGTGAACGTGGAGGCTTCGAAGTTCA  141
                 ||||||||||||||||  |||||||||||||||||||||||||||  ||||||||||||||
Sbjct  4182828  TCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCG  4182887

Query  142      AGACGTACACCCGACTCACTACGGCCGCTTATGTACAATCGAAACCCCAGAAGCA-CGAA  200
                |||||||||||||||||||||||||  |||  |||||  ||||||||||||||  ||||    ||||
Sbjct  4182888  AGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAA  4182947

Query  201      CATCGGACTGATCAACTCTCTGTCCGTG  228
                ||||||  ||||||||||||||||||||
Sbjct  4182948  CATCGGTCTGATCAACTCTCTGTCCGTG  4182975
```


2016/4 BF II
Sequence: 29413690
Samples: 15701
Bases: 245
Average spacing: 65.0
Average quality >= 10: 152, 20: 39, 30: 4
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
16.05.2019
G G A G A T T G T C G GG T C A G A T A T G C TGCA G T C A GC T G T C T A G T T T A T G G C CA G A A G A C C C C GC T G T C T T G A G A T T A C G T A G A A C
G A C G T A T C T C C G C A C T C G A C C C A G G C G G T C T G A C C C G T G A A C G T GG A G GC T T C G A A G T T C A A G A C G T A C A C C C G A C T C A C T A C G G C
C G C T T A T G T A C A A T C G A A A C C C C C A G A A G C A C G A A C A T C G G A C T G A T C A A C T C T C T G T C C G T G G C A C A A G A A A A A A A A A C

Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182989 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
344 bits(186)	2e-95	212/224(95%)	3/224(1%)	Plus/Plus

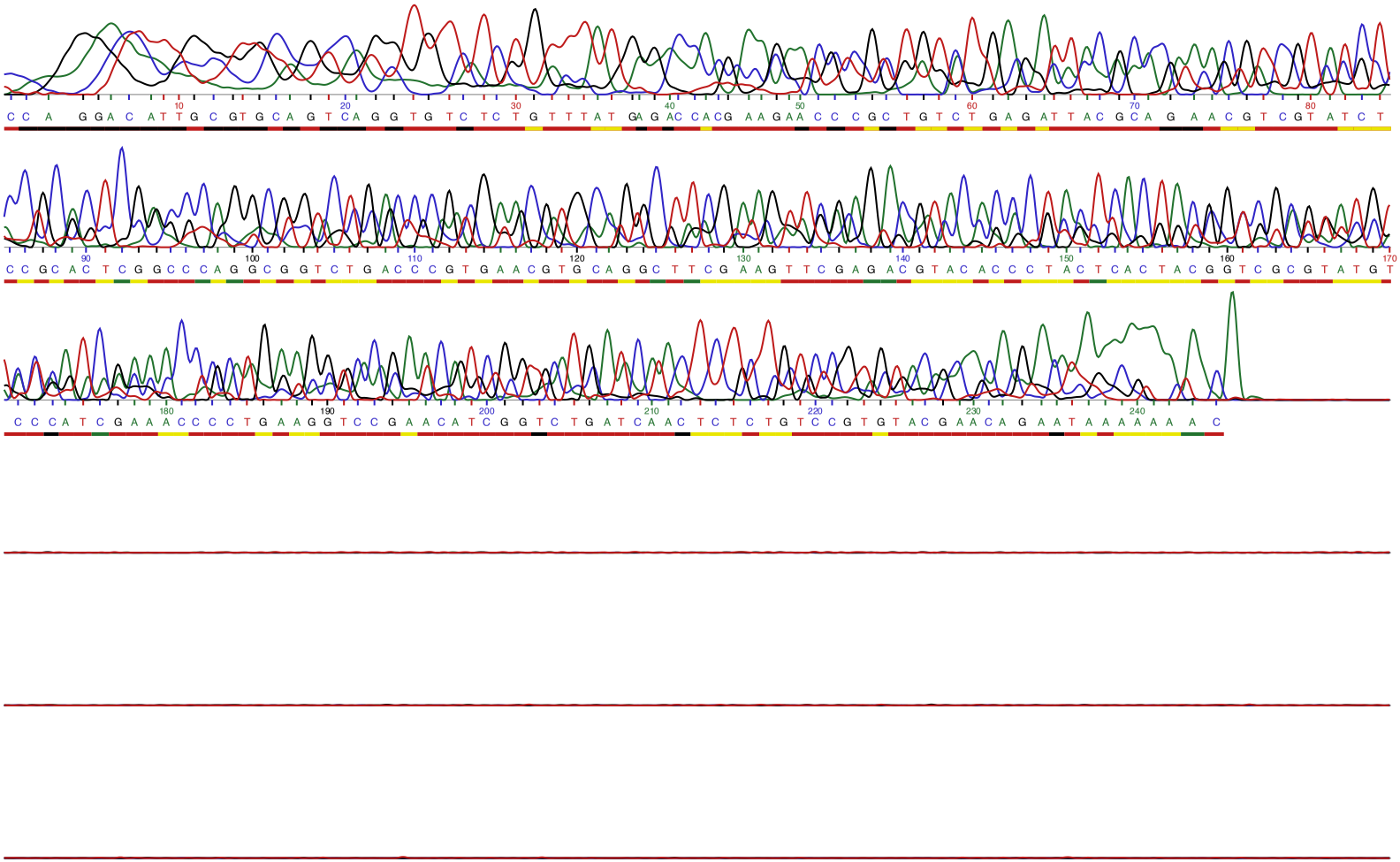
```
Query  16       CAGTCAGGTGTCTCTGTTTATGAGACCACGAAGAACCCGCTGTCTGAGATTACGCA-GAA  74
                |||  |||  ||||||  |||||||  |||||  |||  ||||||||||||||||||||||||   ||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCA-GAACAACCCGCTGTCTGAGATTACGCACAAA  4182825

Query  75       CGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182826  CGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  4182885

Query  135      CGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCCATCGAAACCCCTGAAGGTCCG  194
                |||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  4182886  CGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCG  4182945

Query  195      AACATCGGTCTGATCAACTCTCTGTCCGTGTACGAACAGAATaa  238
                |||||||||||||||||||||||||||||||||||| |||||  |||
Sbjct  4182946  AACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2016/4 BF III

Sequence: 29413997

Samples:	15284	Quality: 0 - 9	
Bases:	245	10 - 19	
Average spacing:	63.0	20 - 29	
Average quality >=	10: 132, 20: 73, 30: 10	>= 30	

Page: 1 / 3
16.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182990 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
364 bits(197)	1e-101	216/225(96%)	2/225(0%)	Plus/Plus

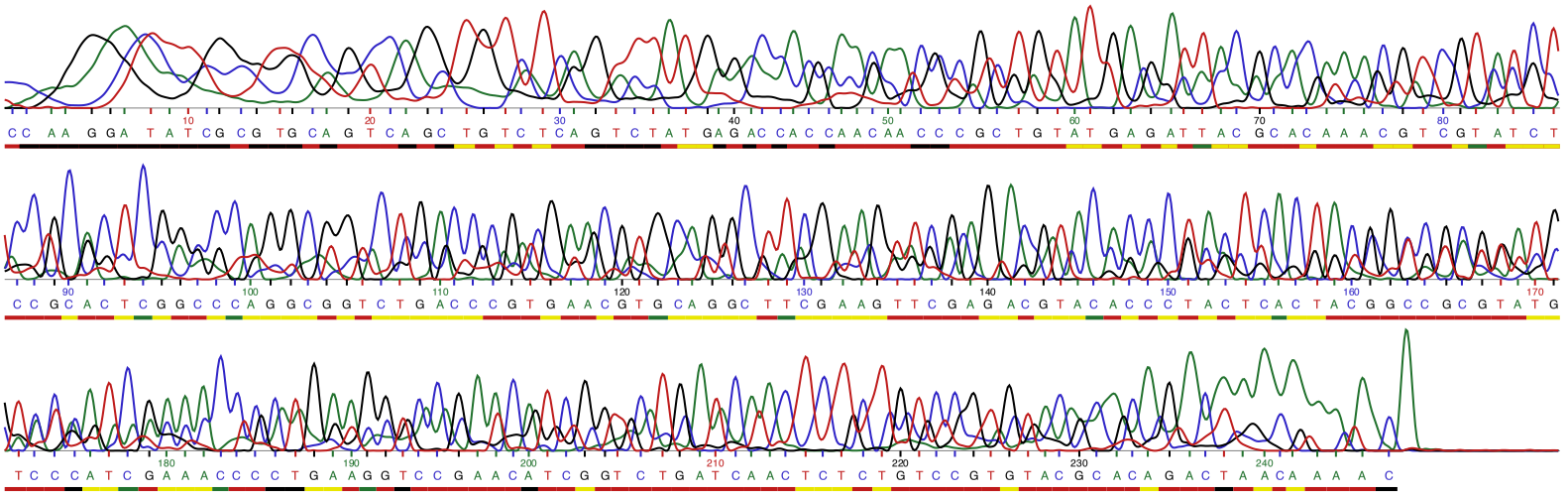
```
Query  17       CAGTCAGCTGTCTCAGTCTATGAGACCACCAACAACCCGCTGTATGAGATTACGCACAAA  76
                |||  |||||||||||||| ||||  ||||||  |||||||||||||| ||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCA-GAACAACCCGCTGTCTGAGATTACGCACAAA  4182825

Query  77       CGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182826  CGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTT  4182885

Query  137      CGAGACGTACACCCTACTCACTACGGCCGCGTATGTCCCATCGAAACCCCTGAAGGTCCG  196
                |||||||||||||| |||||||||||  |||||||||||  ||||||||||||||||||||
Sbjct  4182886  CGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCG  4182945

Query  197      AACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAAC  241
                |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182946  AACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAAC  4182990
```

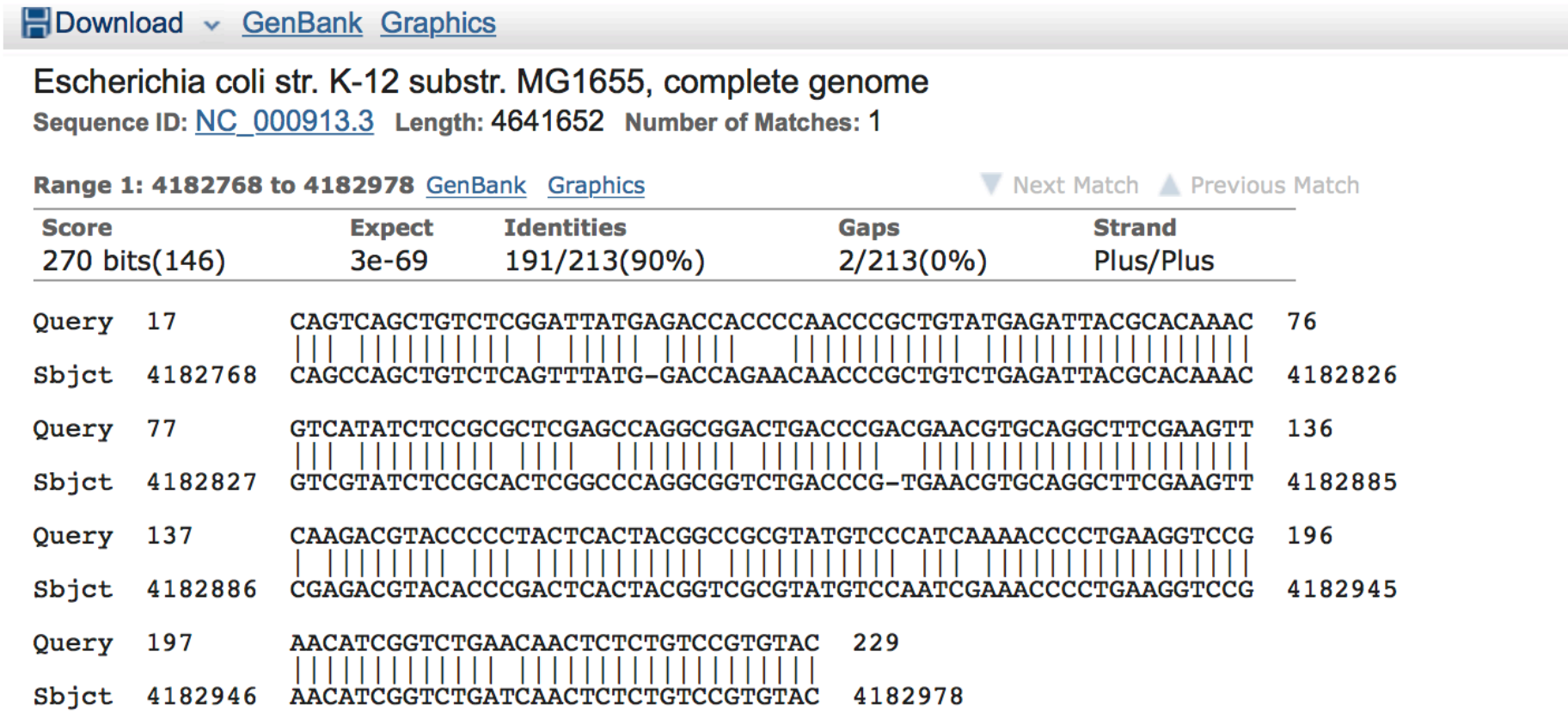
2016/4 BF V

Sequence: 29414000

Samples:	15226	Quality:	0 - 9
Bases:	247		10 - 19
Average spacing:	62.0		20 - 29
Average quality >=	10: 127, 20: 75, 30: 11		>= 30

Page: 1 / 3
16.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182978 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
270 bits(146)	3e-69	191/213(90%)	2/213(0%)	Plus/Plus

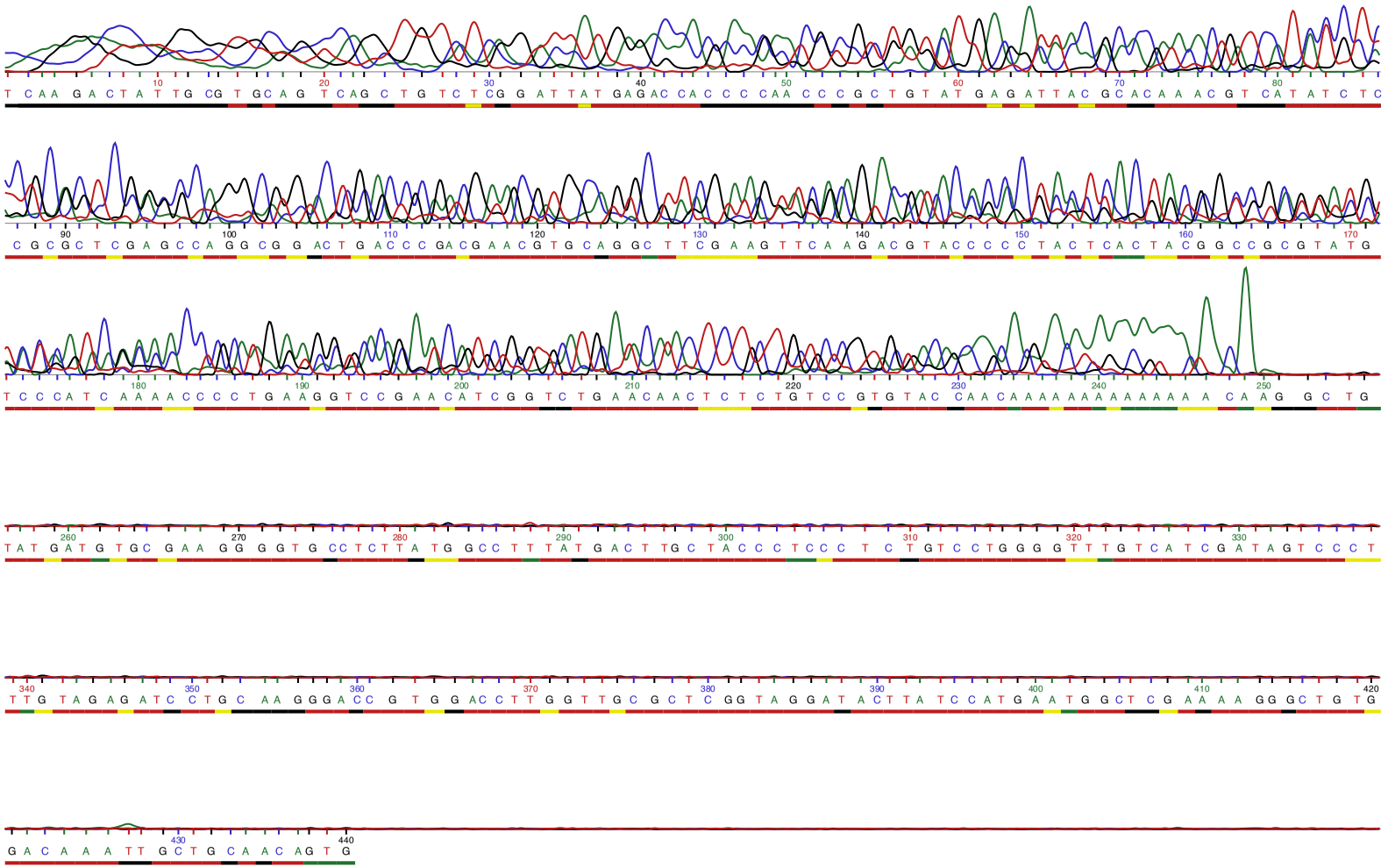
```
Query  17       CAGTCAGCTGTCTCGGATTATGAGACCACCCCAACCCGCTGTATGAGATTACGCACAAAC  76
                |||  |||||||||| | |||||  |||||   |||||||||||  ||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  77       GTCATATCTCCGCGCTCGAGCCAGGCGGACTGACCCGACGAACGTGCAGGCTTCGAAGTT  136
                |||  ||||||||| ||||  ||||||||  |||||||  |||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCG-TGAACGTGCAGGCTTCGAAGTT  4182885

Query  137      CAAGACGTACCCCCTACTCACTACGGCCGCGTATGTCCCATCAAAACCCCTGAAGGTCCG  196
                | ||||||||  |||  ||||||||||  ||||||||||  |||  ||||||||||||||||||
Sbjct  4182886  CGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCG  4182945

Query  197      AACATCGGTCTGAACAACTCTCTGTCCGTGTAC  229
                |||||||||||||  ||||||||||||||||||
Sbjct  4182946  AACATCGGTCTGATCAACTCTCTGTCCGTGTAC  4182978
```

2019/19 AF II

Sequence: 29416004

Samples:	16303	Quality:	0 - 9
Bases:	441		10 - 19
Average spacing:	37.0		20 - 29
Average quality >=	10: 296, 20: 61, 30: 21		>= 30

28.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182989 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
355 bits(192)	7e-95	211/220(96%)	1/220(0%)	Plus/Plus

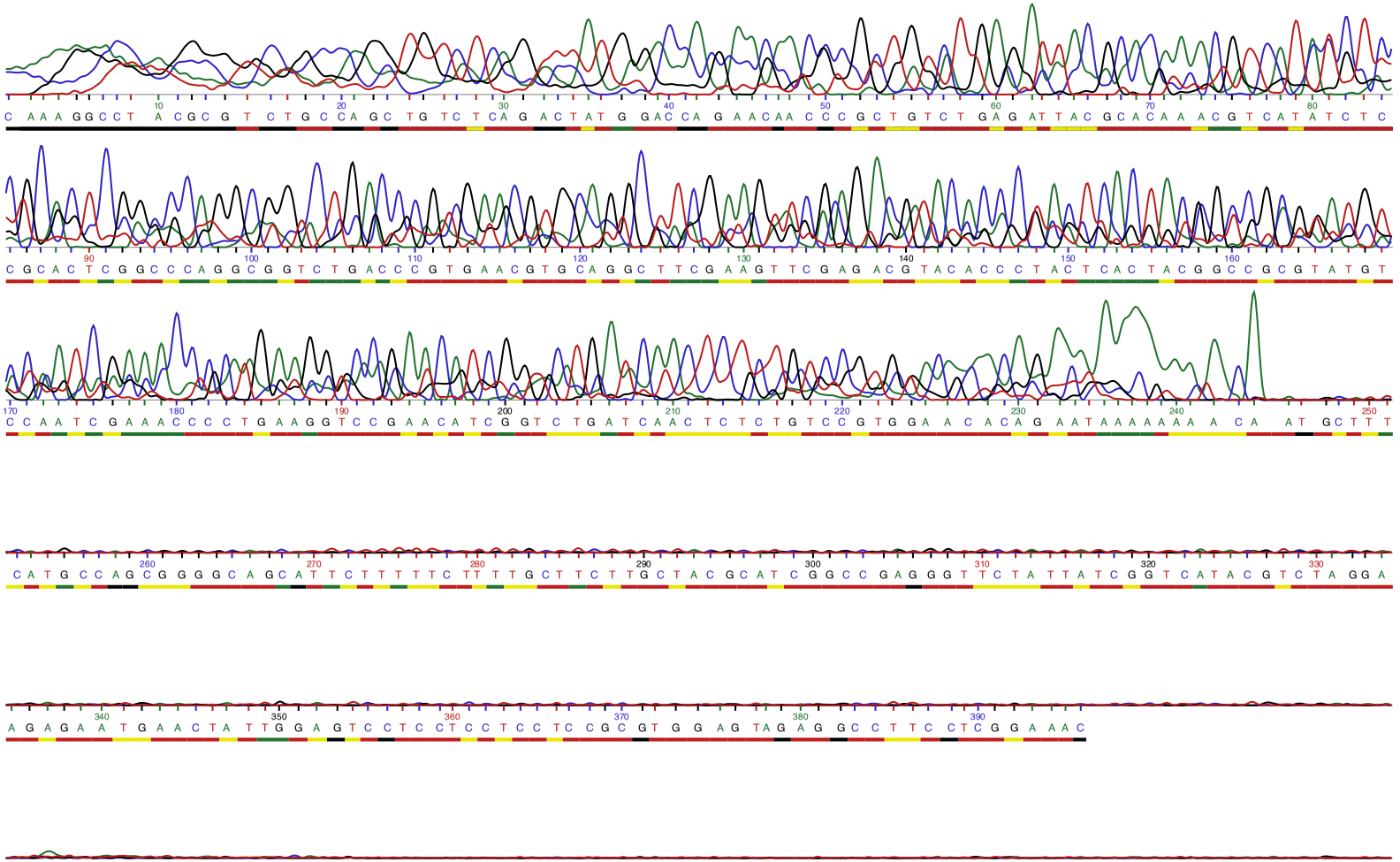
```
Query  18       GCCAGCTGTCTCAGACTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  77
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  78       ATATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  137
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  138      ACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  197
                ||||||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  198      TCGGTCTGATCAACTCTCTGTCCGTGGAA-CACAGAATaa  236
                |||||||||||||||||||||||||||| |  ||||||| |||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

2019/19 BF II

Sequence: 29416080

Samples:	16221	Quality:	0 - 9
Bases:	397		10 - 19
Average spacing:	41.0		20 - 29
Average quality >=	10: 213, 20: 99, 30: 47		>= 30

28.05.2019



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182770 to 4182989 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
355 bits(192)	1e-98	211/220(96%)	1/220(0%)	Plus/Plus

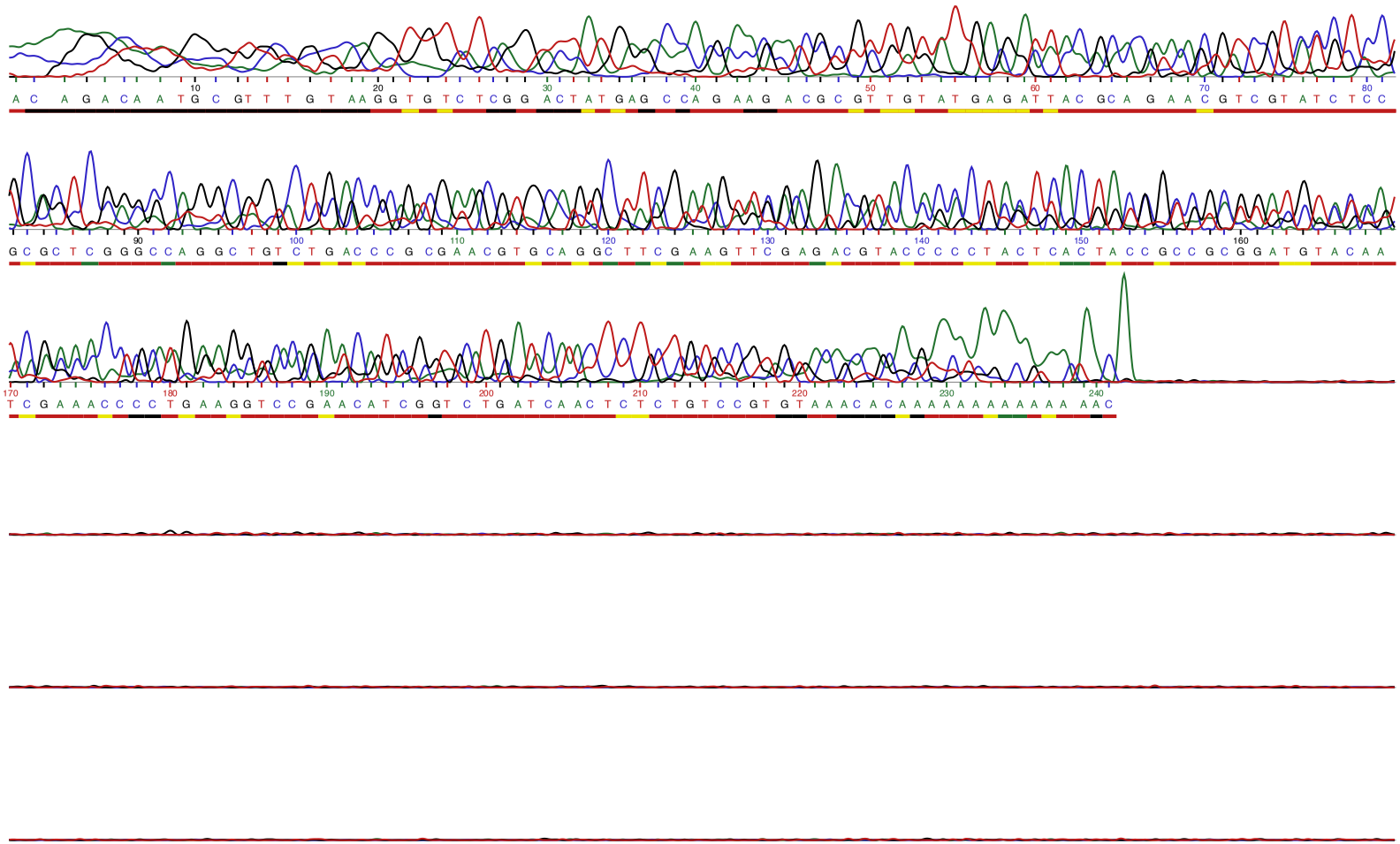
```
Query  18       GCCAGCTGTCTCAGACTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  77
               ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182770  GCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTC  4182829

Query  78       ATATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  137
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182830  GTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAG  4182889

Query  138      ACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  197
               |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4182890  ACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACA  4182949

Query  198      TCGGTCTGATCAACTCTCTGTCCGTGGAA-CACAGAATaa  236
               |||||||||||||||||||||||||||| |  ||||||| |||
Sbjct  4182950  TCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAA  4182989
```

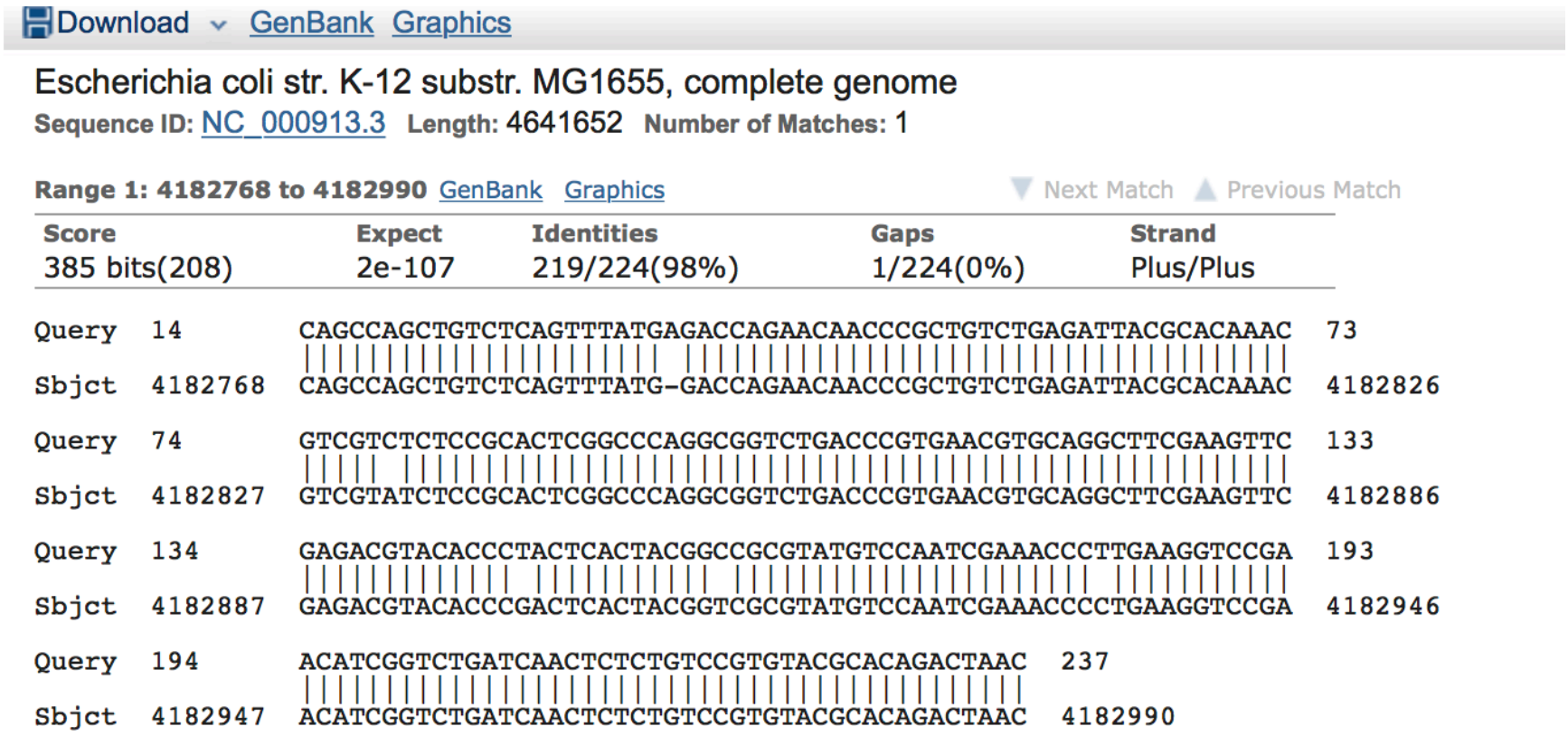
2019/19 DF II

Sequence: 29416097

Samples:	15719	Quality:	0 - 9
Bases:	242		10 - 19
Average spacing:	65.0		20 - 29
Average quality >=	10: 149, 20: 44, 30: 10		>= 30

28.05.2019





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182990 GenBank Graphics Next Match Previous Match

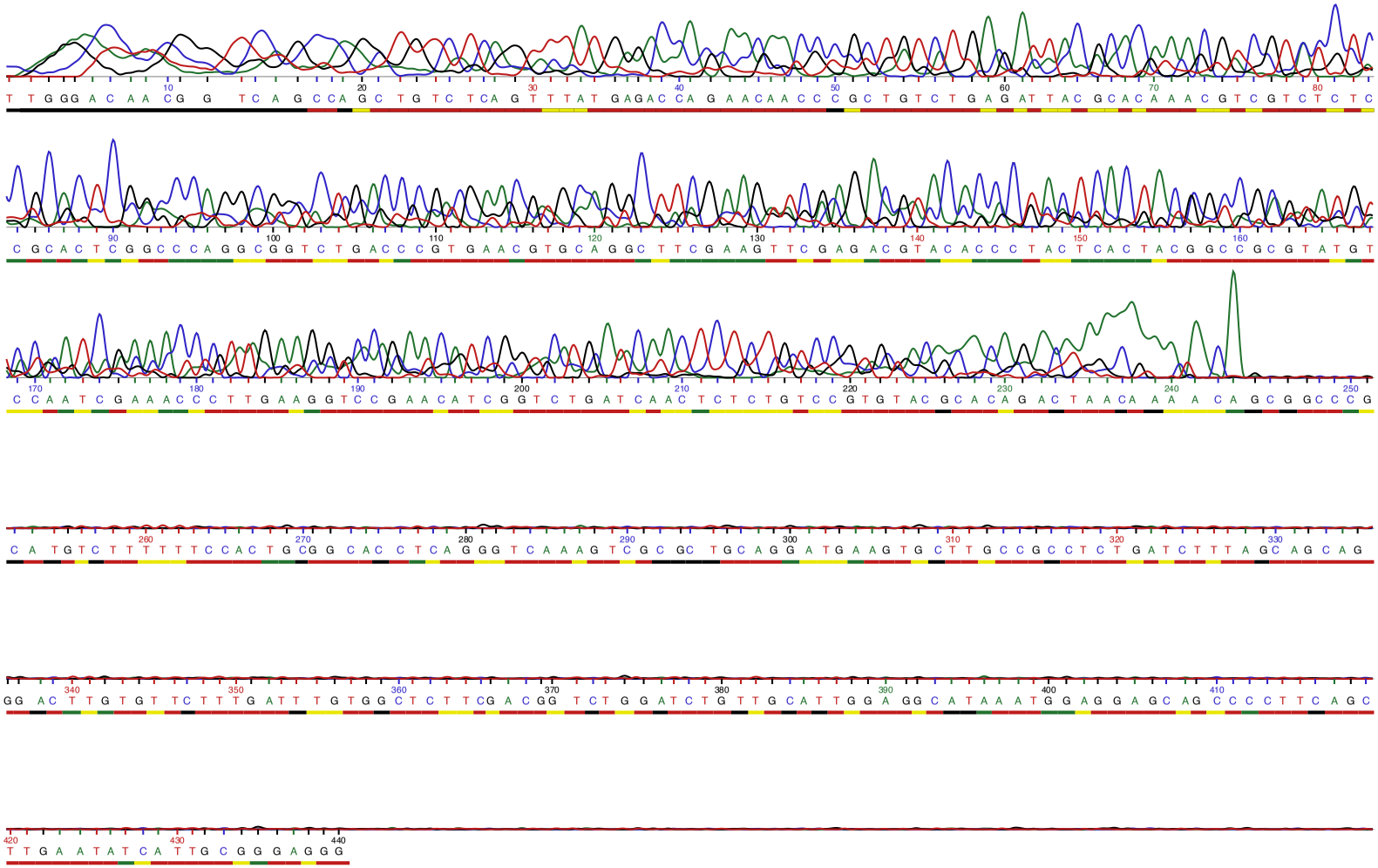
Score	Expect	Identities	Gaps	Strand
385 bits(208)	2e-107	219/224(98%)	1/224(0%)	Plus/Plus

```
Query  14       CAGCCAGCTGTCTCAGTTTATGAGACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  73
                |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  74       GTCGTCTCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  133
                |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  134      GAGACGTACACCCTACTCACTACGGCCGCGTATGTCCAATCGAAACCCTTGAAGGTCCGA  193
                |||||||||||||  ||||||||||  |||||||||||||||||||||  ||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  194      ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAAC  237
                ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAAC  4182990
```

2019/21 AK II

Sequence: 29415328

Samples:	16300	Quality:	0 - 9
Bases:	441		10 - 19
Average spacing:	37.0		20 - 29
Average quality >=	10: 250, 20: 94, 30: 49		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182760 to 4182974 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
357 bits(193)	4e-99	209/216(97%)	3/216(1%)	Plus/Plus

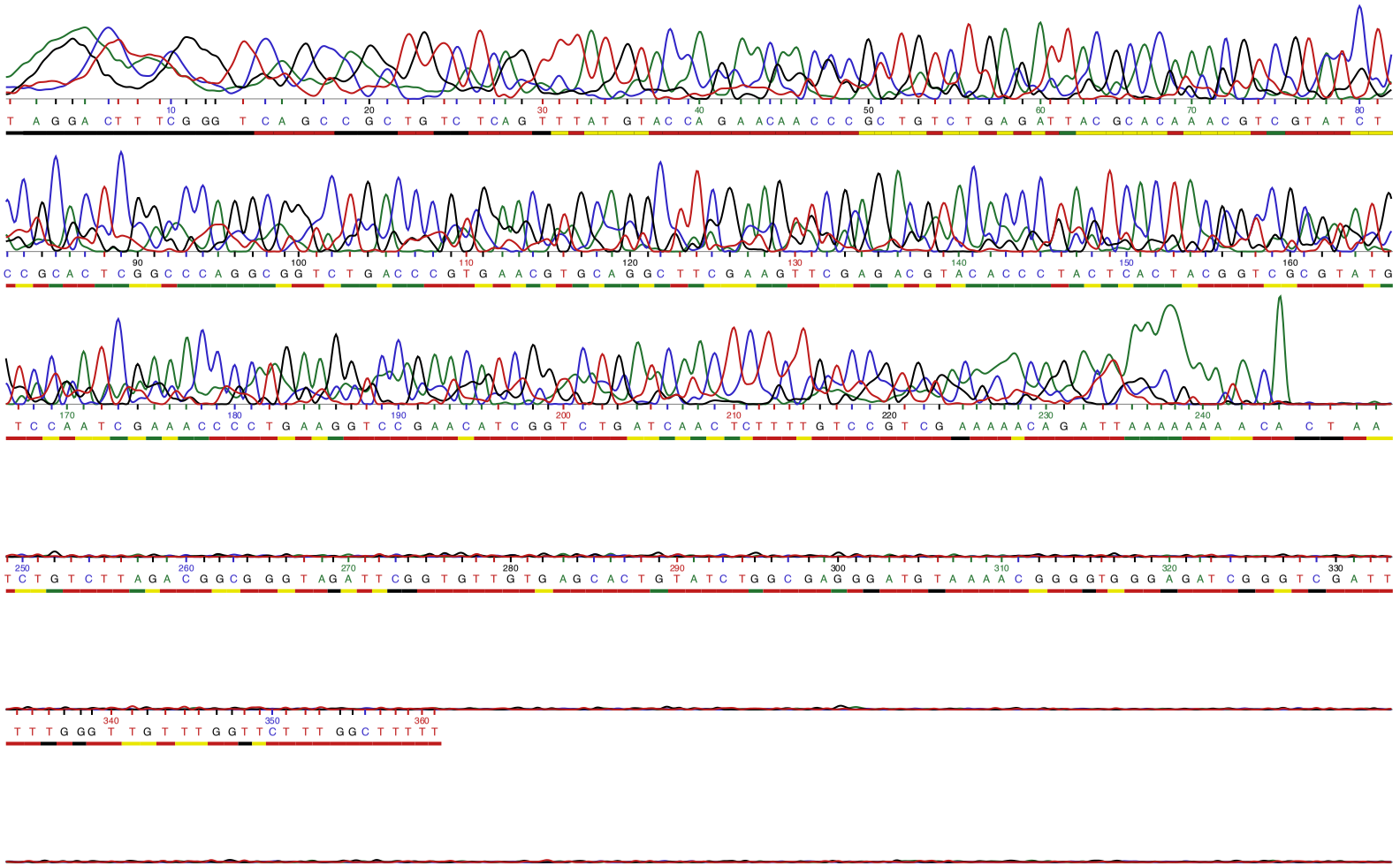
```
Query  8        TTCGGGT-CAGCC-GCTGTCTCAGTTTATGTACCAGAACAACCCGCTGTCTGAGATTACG  65
               ||||| | ||||| ||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  4182760  TTCGGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACG  4182819

Query  66       CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182820  CACAAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTC  4182879

Query  126      GAAGTTCGAGACGTACACCCTACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  185
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  4182880  GAAGTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAA  4182939

Query  186      GGTCCGAACATCGGTCTGATCAACTCTTTTGTCCGT  221
               |||||||||||||||||||||||||| | ||||||
Sbjct  4182940  GGTCCGAACATCGGTCTGATCAACTCTCT-GTCCGT  4182974
```

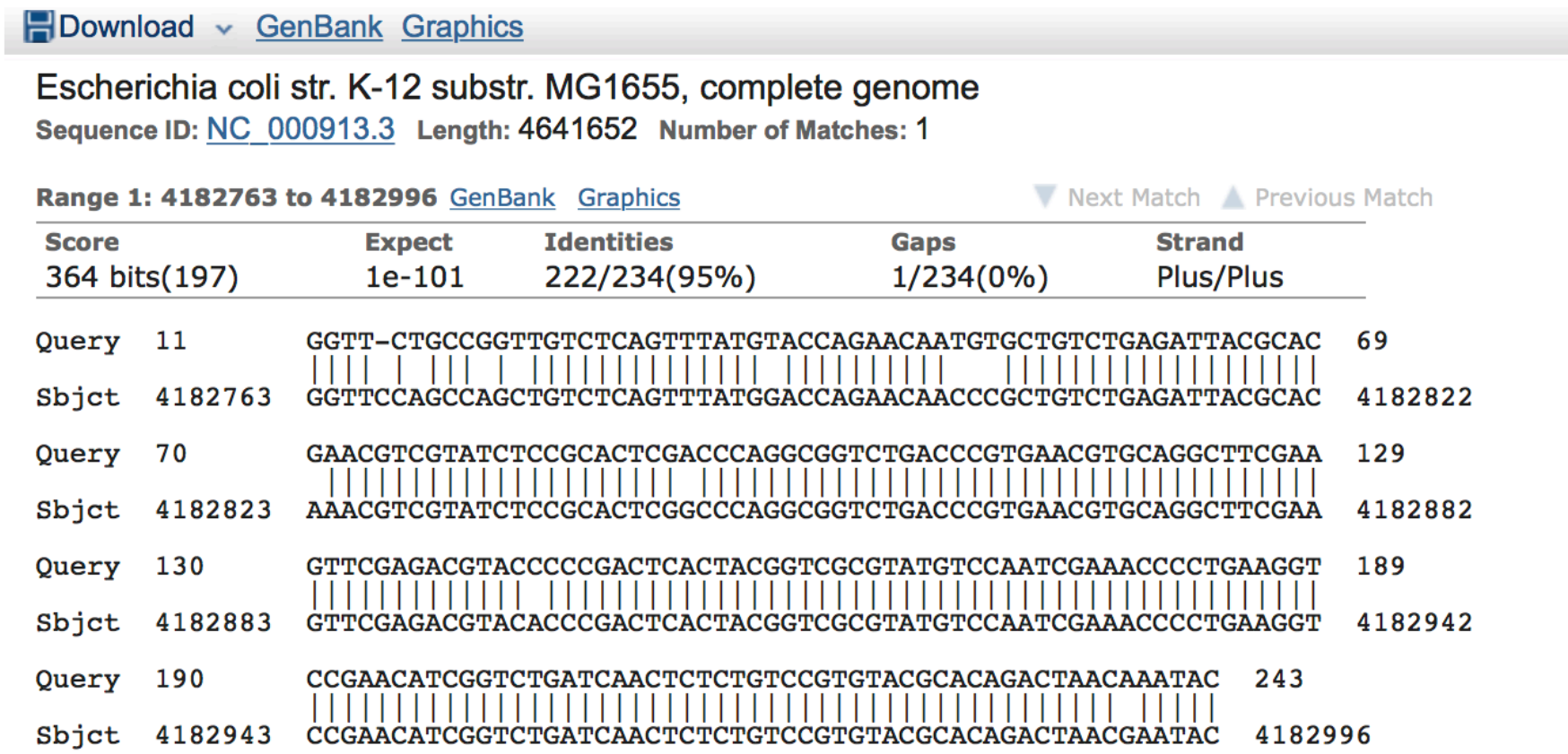
2019/21 AF

Sequence: 29415427

Samples:	16300	Quality: 0 - 9	
Bases:	362	10 - 19	
Average spacing:	46.0	20 - 29	
Average quality >=	10: 190, 20: 85, 30: 52	>= 30	

22.05.2019

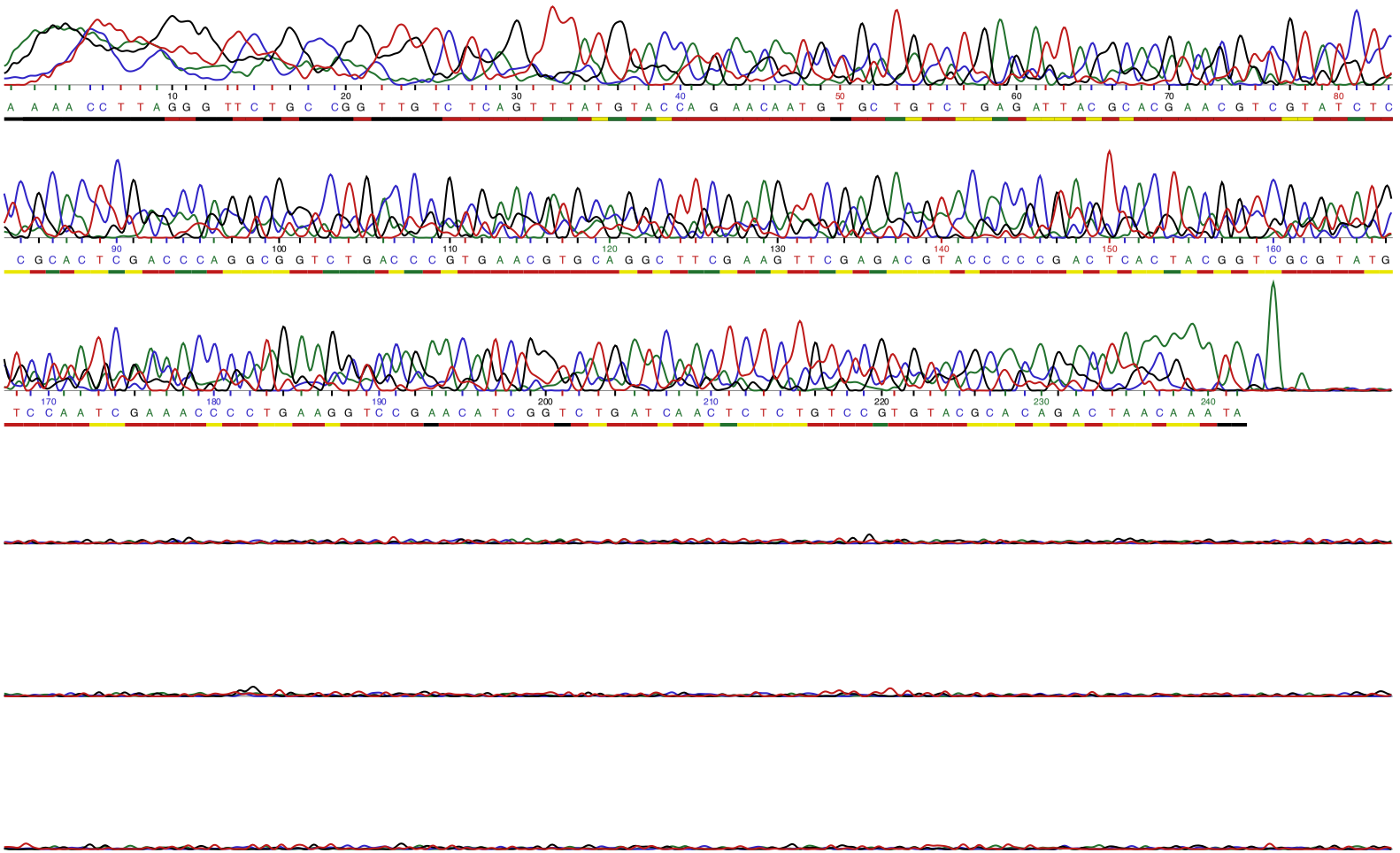



Download GenBank Graphics
Escherichia coli str. K-12 substr. MG1655, complete genome
Sequence ID: NC_000913.3 Length: 4641652 Number of Matches: 1
Range 1: 4182763 to 4182996 GenBank Graphics
Next Match Previous Match
Score 364 bits(197)
Expect 1e-101
Identities 222/234(95%)
Gaps 1/234(0%)
Strand Plus/Plus
Query 11 GGTT-CTGCCGGTTGTCTCAGTTTATGTACCAGAACAATGTGCTGTCTGAGATTACGCAC 69
Sbjct 4182763 GGTTCCAGCCAGCTGTCTCAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCAC 4182822
Query 70 GAACGTCGTATCTCCGCACTCGACCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAA 129
Sbjct 4182823 AAACGTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAA 4182882
Query 130 GTTCGAGACGTACCCCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGT 189
Sbjct 4182883 GTTCGAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGT 4182942
Query 190 CCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACAAATAC 243
Sbjct 4182943 CCGAACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGACTAACGAATAC 4182996

2019/21 CF

Sequence: 29415458

Samples:	16303	Quality:	0 - 9
Bases:	243		10 - 19
Average spacing:	68.0		20 - 29
Average quality >=	10: 130, 20: 65, 30: 23		>= 30



Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182772 to 4182977 GenBank Graphics Next Match Previous Match

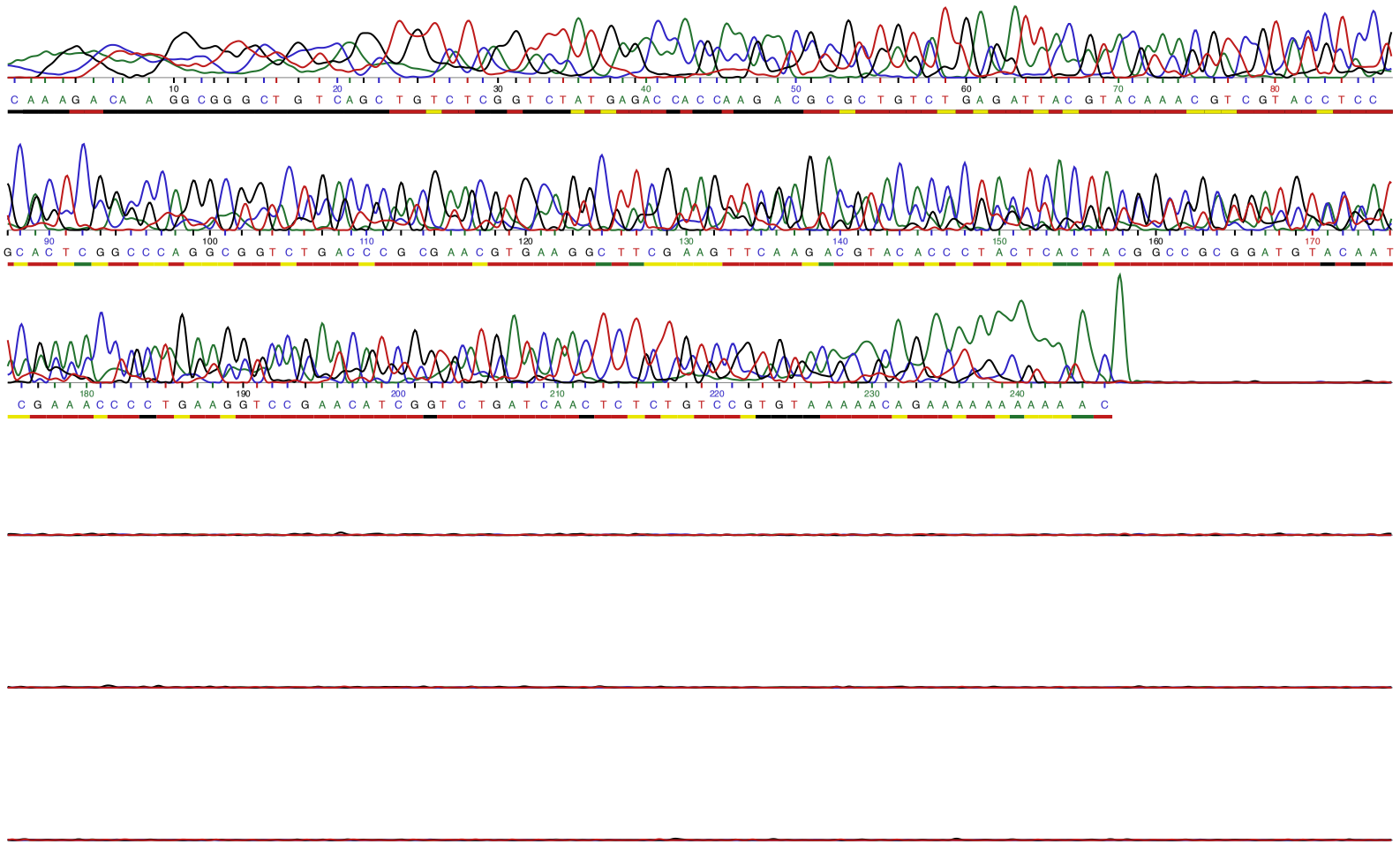
Score	Expect	Identities	Gaps	Strand
287 bits(155)	3e-78	190/207(92%)	1/207(0%)	Plus/Plus

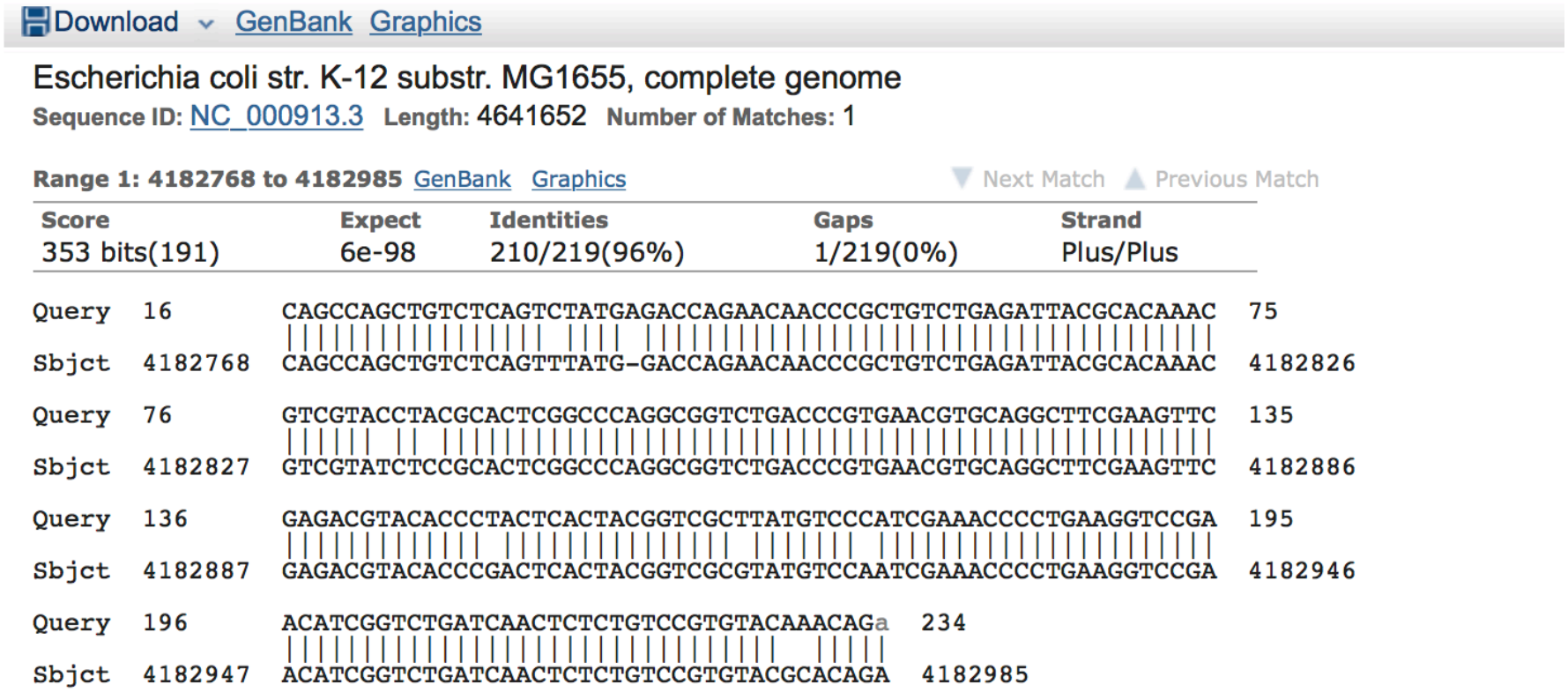
```
Query  20       CAGCTGTCTCGGTCTATGAGACCACCAAGACGCGCTGTCTGAGATTACGTACAAACGTCG  79
                ||||||||||| || |||| |||||  |  || ||||||||||||||||| ||||||||||
Sbjct  4182772  CAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCG  4182830

Query  80       TACCTCCGCACTCGGCCCAGGCGGTCTGACCCGCGAACGTGAAGGCTTCGAAGTTCAAGA  139
                || ||||||||||||||||||||||||||||| ||||||| ||||||||||||| |||
Sbjct  4182831  TATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGA  4182890

Query  140      CGTACACCCTACTCACTACGGCCGCGGATGTACAATCGAAACCCCTGAAGGTCCGAACAT  199
                ||||||||| |||||||||| |||| |||| ||||||||||||||||||||||||||||
Sbjct  4182891  CGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACAT  4182950

Query  200      CGGTCTGATCAACTCTCTGTCCGTGTA  226
                |||||||||||||||||||||||||||
Sbjct  4182951  CGGTCTGATCAACTCTCTGTCCGTGTA  4182977
```

2019/21 DK II

Sequence: 29416110

Samples:	15704	Quality: 0 - 9	
Bases:	246	10 - 19	
Average spacing:	64.0	20 - 29	
Average quality >=	10: 145, 20: 51, 30: 8	>= 30	





Download GenBank Graphics

Escherichia coli str. K-12 substr. MG1655, complete genome

Sequence ID: NC_000913.3 **Length:** 4641652 **Number of Matches:** 1

Range 1: 4182768 to 4182985 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
353 bits(191)	6e-98	210/219(96%)	1/219(0%)	Plus/Plus

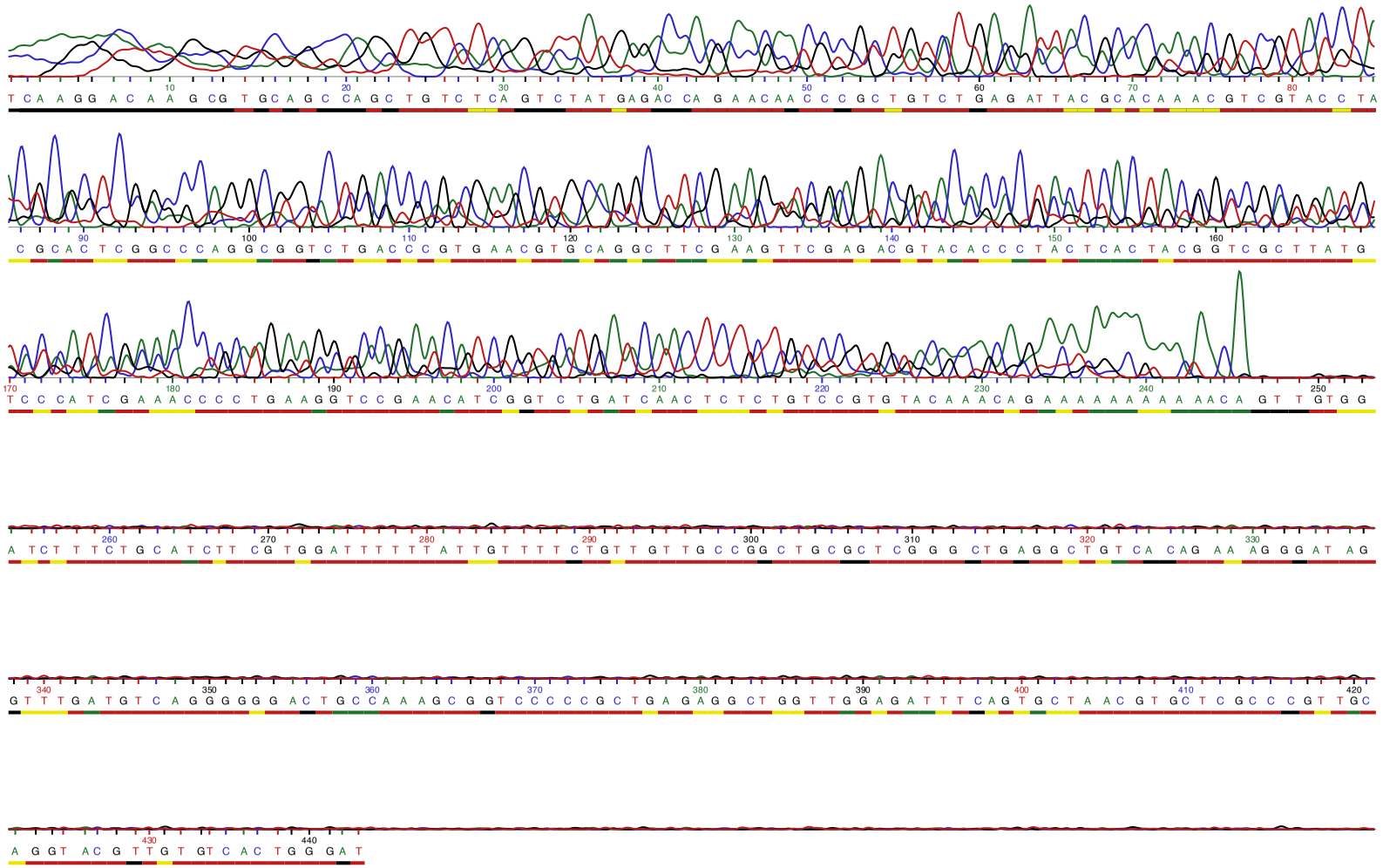
```
Query  16       CAGCCAGCTGTCTCAGTCTATGAGACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  75
                ||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4182768  CAGCCAGCTGTCTCAGTTTATG-GACCAGAACAACCCGCTGTCTGAGATTACGCACAAAC  4182826

Query  76       GTCGTACCTACGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  135
                |||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4182827  GTCGTATCTCCGCACTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTC  4182886

Query  136      GAGACGTACACCCTACTCACTACGGTCGCTTATGTCCCATCGAAACCCCTGAAGGTCCGA  195
                |||||||||||||| ||||||||||||||| |||||||| ||||||||||||||||||||
Sbjct  4182887  GAGACGTACACCCGACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGA  4182946

Query  196      ACATCGGTCTGATCAACTCTCTGTCCGTGTACAAACAGa  234
                |||||||||||||||||||||||||||||||||||||  |||||
Sbjct  4182947  ACATCGGTCTGATCAACTCTCTGTCCGTGTACGCACAGA  4182985
```

2019/21 DF

Sequence: 29416127

Samples:	16303	Quality: 0 - 9	
Bases:	444	10 - 19	
Average spacing:	37.0	20 - 29	
Average quality >=	10: 270, 20: 85, 30: 40	>= 30	

28.05.2019



A.2 DNA Glycosylase activities for N^4,5-dimethylcytosine

A.2.1 Buffers and solutions

Table A.2.1: 5× HEPES buffer

Composition	Stock	Amount used in preparation
225 mM HEPES, pH 7.5	1 M (lab stock)	22,5 ml
10% glycerol	85% (Merck, Cat # 1.04094)	12 ml
2 mM EDTA	0.5 M (lab stock)	400 µl
Deionized H_2O		Dilute to 100 µl

Stored in aliquots at -20°C

Table A.2.2: Salt-TE (STE) buffer

Composition	Stock	Amount used in preparation
10 mM Tris, pH 8.0	1 M	200 µl
50 mM NaCl	Sigma, Cat. #S5886, 58.44 g/mol	58.44 mg
1 mM EDTA	0.5 M (lab stock)	40 µl
Deionized H_2O		Dilute to 20 µl

Filtrated and stored in aliquots at -20°C

Table A.2.3: 1× TE buffer

Composition	Stock	Amount used in preparation
10 mM Tris, pH 7.5	1 M	200 µl
1 mM EDTA, pH 8.0	0.5 M (lab stock)	1.6 ml
Deionized H_2O		Dilute to 20 ml

Filtrated and stored in aliquots at -20°C

Table A.2.4: 10X TBE running buffer

Composition	Stock	Amount used in preparation
890 mM Tris base	Sigma, Cat. # T6066, 121,14 g/mol	108 g
890 mM Boric acid	Mw 61.8 g/mol	55 g
20 mM EDTA, pH 8	0.5 M (lab stock)	40 ml
Deionized H_2O		Dilute to 1000 ml

Autoclave before use

Table A.2.5: Loading buffer

Composition	Stock	Amount used in preparation
formamide, 80%	99.5% (Sigma, Cat. # F9037)	40.20 ml
1 mM EDTA	0.5 M (lab stock)	100 µl
Blue dextran, 1% (w/v)	Sigma, Cat. #D5751	0.5 g
Deionized H_2O		Dilute to 50 ml

Stored in aliquots at -20°C

Table A.2.6: 96% ethanol with 0.1 M CH_3COONa

Composition	Stock	Amount used in preparation
0.1 M NaOAc	Sigma, Cat. # S2889, 82.03 g/mol	0.82 g
96% EtOH	99.5%	225 ml

Stored in RT, cooled down to -20°C before use

Table A.2.7: 1 M KCl

Composition	Stock	Amount used in preparation
1 M KCl	Sigma, Cat. # S5405, 74.55 g/mol	3.72 g
Deionized H_2O		Dilute to 50 ml

Filtrated and stored in aliquots at -20°C

Table A.2.8: Stop solution

Composition	Stock	Amount used in preparation
20 mM EDTA	0.5 M (lab stock)	2 ml
Sodium dodecyl sulphate (SDS), 0.5% (w/v)	99% Sigma, Cat. #L3771, 288.38 g/mol	252 mg
Deionized H_2O		Dilute to 50 ml

Stored at room temperature

Table A.2.9: Denaturing 20% PAGE gel with 8 M urea

Composition	Stock	Amount used in preparation
Polyacrylamide, 20% (w/v)	40% (Saveen Werner AB, Cat. #BIAC21)	3.5 ml
1× TBE	10× TBE (lab stock)	700 µl
urea	99.5%, Sigma, Cat. #F9037	3.363 g
Deionized H_2O		280 µl
ammonium persulfate (APS)	BioRad, Cat. #161-0700, 228.2 g/mol)	35 µl
tetramethylethylenediamine (Temed)	Invitrogen Cat. #15524-010	3.5 µl

A.2.2 Detailed protocols

A.2.2.1 Hybridization of oligos

Tubes kept on ice and in darkness during assay.

1. Prepare reaction mixture in PCR tube: 2 µl forward ssDNA (100 pmol/µl)
   2 µl of complementary strand ssDNA (100 pmol/ml)
   16 µl 1× Salt-Tris-EDTA (STE) buffer
2. Incubate at 95°C for 4 min in the thermocycler.
3. Leave the tube in the thermocycler to cool down at 1°C per min for 2 hours.
4. Dilute with 180 µl 1× Tris-EDTA (TE) buffer to 1 pmol/µl. Store in aliquots at -20°C in the dark.

A.2.2.2 Base excision assay

Whole assay is performed on ice and in darkness.

1. Prepare reaction mix according to table A.2.10, note that DTT has to be made fresh for every assay.

Table A.2.10: Reaction mixtures for base excision assays

	Reagent	Stock	Reaction 1x
hSMUG (25-270)	HEPES buffer with 5 mM DTT	5×	4 µl
	KCl	1 M	1.4 µl
	BSA	10 mg/ml	1 µl
	Deionized H_2O		11.6 µl
	Substrate	1 pmol/µl	1 µl
	Enzyme, hSMUG (25-270)		1 µl/0µl*
	HEPES buffer with 5mM DTT	1×	1 µl/0µl*
	Total volume		20 µl
MPG	ThermoPol buffer		2 µl
	DTT	20 mM	1 µl
	BSA	10 mg/ml	1 µl
	Deionized H_2O		14 µl
	Substrate	1 pmol/µl	1 µl
	Enzyme, MPG		1 µl/0 µl
	ThermoPol buffer		1 µl/0 µl
	Total volume		20 µl
MutY	REC buffer		2 µl
	DTT	20 mM	1 µl
	BSA	10 mg/ml	1 µl
	Deionized H_2O		14 µl
	Substrate	1 pmol/µl	1 µl
	Enzyme, MutY		1 µl/1 µl
	REC buffer		1 µl/1 µl
	Total volume		20 µl

*For negative controls 1 µl of buffer solution is added instead of enzyme to achieve equal volume in all tubes.

2. Centrifuge 1min, 4000 rpm at RT.

3. Incubate at 37°C for 1 hour (water bath)

4. Spin down and put on ice

5. Add 45 μl Stop solution and 1 μl Proteinase K

6. Spin down

7. Incubate at 37°C for 10 min (water bath)

8. Spin down and put on ice

9. Add 150 μl 96% EtOH w/0.1 M NaAc (cold from freezer)

10. Add 1.6 μl tRNA (10 mg/ml), and invert the tubes several times

11. Incubate the tubes at -70°C, 2 h, or -20°C overnight in darkness.

12. Centrifuge the tubes; 13 000 rpm, 15 min, 4°C

13. Rotate the tubes 180°

14. Centrifuge again; 13 000 rpm, 15 min, 4°C

15. Remove supernatant

16. Add 300 μl cold 70% EtOH (from freezer)

17. Centrifuge tubes; 13 000 rpm, 5 min, 4°C

18. Remove supernatant

19. Centrifuge the tubes; 13 000 rpm, 1 min, 4°C

20. Remove remaining supernatant by pipette

21. Leave pellet to dry in hood for at least 20 min (with open caps), make sure all EtOH has evaporated.

22. Dissolve pellet in 10 μl 0,1 M NaOH, and incubate; 10 min, 90°C

23. Spin down

24. Add 10 μl denaturing loading buffer (DLB), and incubate; 5 min, 90°C

25. Load 5 μl of each sample to the prewashed wells of the polyacrylamide gel containing 8 M urea.

26. Run gel; 2h, 200 V constant, in darkness

27. Analyse gel

A.2.2.3 Gel preparation, loading and running

1. Assemble the gel cassette as instructed by your supervisor.

2. Check for leakage with dH_2O, the system must not leak, dry with filterpaper upon confirmation of sealed system.

3. Prepare first part of mixture:
   - 3.363 g Urea
   - 3.5 ml 40% acrylamide solution
   - 700 µl 10× TBE buffer
   - 280 µl deionized water (MQ)

4. Microwave for 8-10 sec to dissolve urea.

5. Cool down for about 10 min.

6. Add 35 µl 10% APS and 3.5 µl Temed and swirl the container gently to mix.

7. Use a plastic pipette and transfer the mixture quickly to the space between the glass plates, and insert the comb before the gel starts to polymerize, make sure there are no air bubbles.

8. Leave the gel to dry for 30-40 min.

9. Remove the gel cassette sandwich from the casting stand and frame, and wipe off any overflow with a damp paper towel (to avoid gel particles in the electrophoresis set up).

10. Assemble the gel cassette sandwich into the Electrode assembly with buffer dam, and fill the gel chamber with 1× TBE running buffer, make sure there are no leakages.

11. When no leakage is confirmed, place it in the Mini Tank, and carefully remove the comb (use both hands).

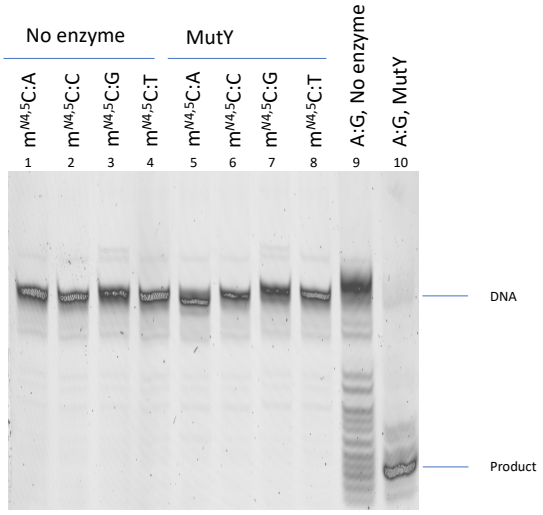
12. Use a 1000 ml pipette to wash the wells, do this several times until you are sure there are no residual gel particles left in the wells (critical to create sharp and even bands).

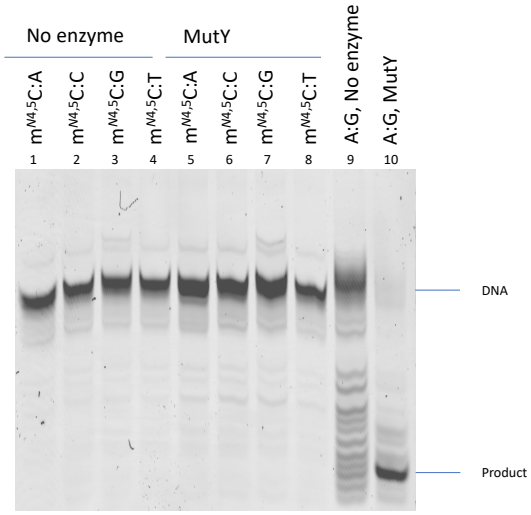
13. Load 5 µl of fully treated samples carefully into the wells.

14. Run the gel at 200 V for 2 h in darkness and RT.

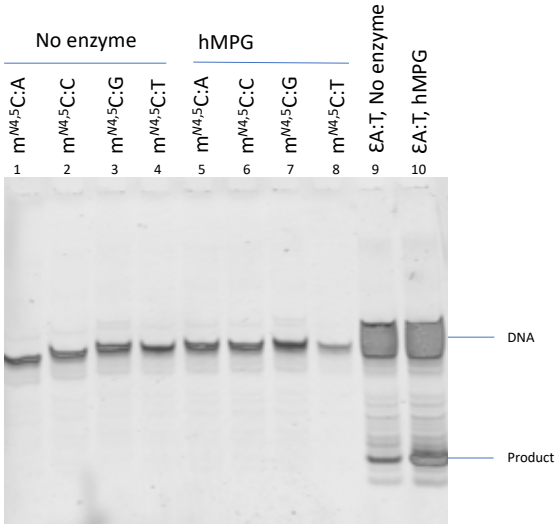
A.2.3 Results, additional gels

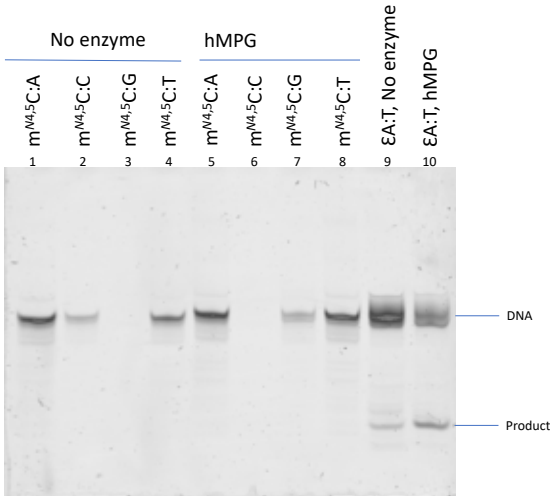
A.2.3.1 MutY



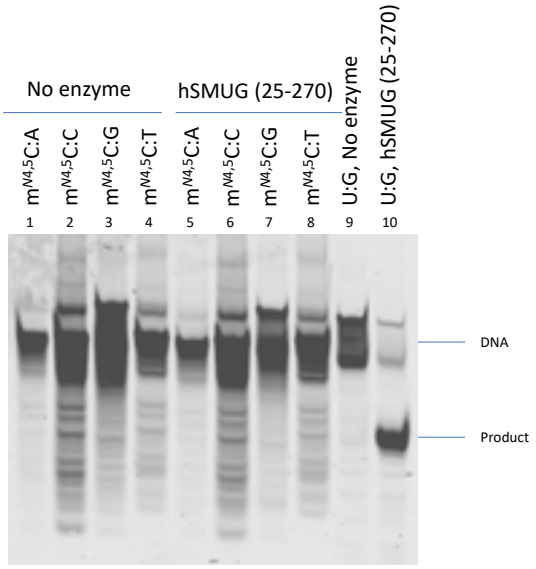


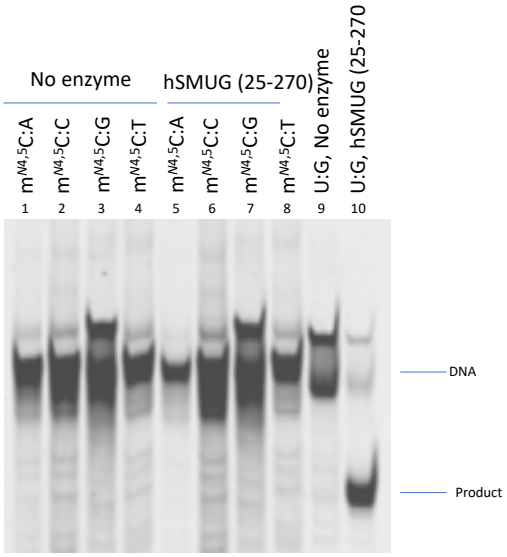
A.2.3.2 hMPG





A.2.3.3 hSMUG (25-270)





A.3 Production and purification of hSMUG (25-270)

A.3.1 Buffers and solutions

Table A.2.11: Buffers and solutions used in production and purification of hSMUG (25-270)

Buffer	Composition
Lysis buffer	50 mM TRIS, pH 7.5, 300 mM NaCl, 5% glycerol
Equilibration buffer	50 mM TRIS, pH 7.5, 300 mM NaCl
Wash buffer	50 mM TRIS, pH 7.5, 300 mM NaCl, 10 mM imidazole
Elution buffer 1	50 mM TRIS, pH 7.5, 300 mM NaCl, 100 mM imidazole
Elution buffer 2	50 mM TRIS, pH 7.5, 300 mM NaCl, 500 mM imidazole
Dialisys buffer 1	50 mM TRIS, pH 7.5, 300 mM NaCl, 2 mM ß-ME
Dialisys buffer 2	50 mM TRIS, ph 7.5, 300 mM NaCl
1× TRIS-Glycine running buffer	25 mM trisma Base, 0,192 M glycine, 1% SDS

A.3.2 Detailed protocols

A.3.2.1 Competent cells

1. Inoculate a single colony of *E.coli* BL21 (DE3) in 3 ml LB media containing and grow overnight (ON) at 37°C with vigorous shaking (220 rpm).

2. Transfer 200 µl of the ON culture to 25 ml of LB media, and grow with vigorous shaking at 37°C until the culture reaches an OD_{600} value of 0.3-0.5.

3. Place the culture on ice for 10 min.

4. Split the culture into round bottom falcon tubes 4×6ml.

5. Centrifuge tubes for 10 min at 4000 rpm and 4°C.

6. Decant and discard supernatant and resuspend in 3 ml ice cold sterile 100 mM $CaCl_2$.

7. Incubate for 30 min on ice.

8. Centrifuge tubes for 10 min at 4000 rpm and 4°C.

9. Decant and discard supernatant and resuspend pellet in 400 µl ice cold $CaCl_2$.

10. Split content of each falcon tube into two Eppendorf tubes (200 µl each).

11. Cells are now competent, and will remain competent for 24 h with decreasing transformation efficiency over time.

12. For storage snap freeze cells in liquid nitrogen and store at -80°C.

A.3.2.2 Transforming bacteria

1. Add 50 ng plasmid to 200 µl aliquot(s) of competent cells from previous step.

2. Incubate on ice for 30 min.

3. Place tube(s) in a 42°C water bath for exactly 90 sec.

4. Place the tube(s) immediately on ice to cool down.

5. Add 1 ml LB media to each tube, and incubate for 1h at 37°C with shaking (225rpm).

6. Plate 200 µl of the culture(s) on LB plate(s) containing 50 µg/ml Kanamycin.

7. Grow plate(s) overnight at 37°C.

A.3.2.3 Autoinduction

1. Inoculate a single transformed colony in 500 ml ZYM-5052 containing 50 µg/ml Kanamycin. Use 2 L baffled Erlenmeyer flasks to allow for enough oxygen, and to create turbulence while shaking.

2. Incubate with shaking (220 rpm) for 24 h at 28°C.

3. Harvest cells through centrifugation at 6000 rpm for 20 min at RT.

A.3.2.4 Affinity purification

Affinity purification using Tallon Beads

1. Add 7ml lysis buffer for each gram of pellet formed in autoinduction, along with 1 tablet of complete EDTA-free protease inhibitor cocktail. (If necessary, the lysate can be frozen at this time at -20°C.)

2. Supplement the bacterial lysate with: lysozyme (final concentration 100 µg/ml), DNAse I (final concentration 5 µg/ml), RNAse A (final concentration5 µg/ml), Tergitol (final concentration 0.5%) and $MgCl_2$ (final concentration 0.5%)

3. Incubate for 30 min at RT, with gentle orbital shaking.

4. Sonicate the lysate on ice; amplitude 30%. Pulse: 10 sec on, 10 sec off, a total of three times.

5. Remove the insoluble debris by centrifugation: 20 000rpm for 40 min at 4°C.
(Prepare 2 ml of Tallon beads during centrifugation of lysate.

6. Wash and equilibrate according to manufacturers instructions.)

7. Decant the supernatant (crude extract) from previous step and place on ice.

8. Incubate crude extract with Tallon beads for 30 min at 4°C.

9. Separate the flow through from the beads by centrifugation; 900×g, 10 min, 4°C)

10. Analyse flow through by SDS-PAGE to check for 6× (His)tag binding to the beads.

11. Wash beads with 10 ml wash buffer for 10min.

12. Centrifuge beads: (700×g, 5 min, 4°C)

13. Elute the beads using 2 ml elution buffer 1 and incubate for 10 min at 4°C.

14. Collect the elution fraction (now containing the protein of interest) by centrifugation for 5 min at 700×g and 4°C.

15. Repeat elution step to make sure all of the protein is released from the beads.

16: Analyze the elution fractions by SDS-PAGE.

Dialysis

1. Prepare 2 L of dialysis buffer. (Dialysis buffer 1 to remove his Tag, or dialisys buffer 2 if no TEV treatment is involved).

2. Prepare the Pre-wetted RC tubing; cut to a fitting length and place one magnetic clamp on the bottom of each tube membrane.

3. Add 1 ml of each elution fraction to separate tube membranes, and use clamp at the top to seal the membrane. (Add 25 µl of AcTEV protease to the fraction(s) where removal of His tag is desired).

4. Incubate with gentle stirring over night at 4°C.
(When fractions both with and without TEV treatment are dialysed together, a second dialysis should be performed with dialysis buffer 2 for the fractions with no TEV treatment to remove traces of ßME.)

Äkta Start Purification System (LPLC)

1. Prepare the Äkta Start by washing pumps and fractionation tube with water and then equilibration buffer. Wash pump B with elution buffer 2.

2. Wash the column with water, and equilibrate it with equilibration buffer, make sure to wash and equilibrate with at least 3 ml liquid each time.

3. Set flow rate to 1 ml/min, and load sample to be analyzed to the column using the sample valve.

4. Collect the flow through (now containing protein of interest without His tag) in 2 ml fractions.

5. Elute the column stepwise with 50 mM, 125 mM, 250 mM, 375 mM and 500 mM elution buffer 2 with collection size 1 ml, and flow rate 1 ml/min.

6. Analyze fractions from flow through by nano drop and BCA.

A.3.2.5 SDS-PAGE analysis

1. Prepare the precast gel in the electrophoresis chamber with 1× Tris-Glycine buffer, make sure there are no leakages.

2. Mix 10 µl of samples with 10 µl 2× Laemmli Sample Buffer, and heat for 5 min at 95°C.

3. Cool samples on ice for two min before loading 15 µl of the samples into the wells.

4. Use a molecular weight standard as a ladder for measuring size of protein.

5. Run gel for 30 min at 220V.

6. Stain the gel by placing on a plate and rinse with H_2O, heat until boiling in microwave oven, and discard the water. Repeat rinsing step with water 2 times to wash out SDS.

7. After removal of water, add 15 ml of stain and heat until boiling in microwave oven, place on orbital shaker for at least 1h.

8. Discard the stain and add water to remove the background. Repeat until the background is clear.

9. Use ChemiDoc to read and photograph the gel.

A.3.2.6 Measurement of protein concentration

Microplate procedure

1. Prepare the standard according to table A.2.12.

Table A.2.12: Preparation of standards for the BCA microplate analysis

Vial	Volume of Diluent (µL)	Volume and source of BSA (µL)	Final BSA concentration [µg/ml]
A	0	300 of Stock	2000
B	125	375 of Stock	1500
C	325	325 of Stock	1000
D	175	175 of vial B	750
E	325	325 of vial C	500
F	325	325 of vial E	250
G	325	325 of vial F	125
H	400	100 of vial G	25
I	400	0	0 = Blank

2. Use the following formula to determine the total volume of WR (working reagent) required:
(#standards + #unknowns) × (# replicates) × (volume of WR per sample) = total volume WR

3. 200 µl of WR is required for each well in the microplate procedure.

4. Pipette 25 µl of each standard and unknown sample into the wells of the microplate (use 3 parallells for each standard and sample).

5. Add 200 µl of WR into each well and cover the plate. Incubate for 30 min at 37°C.

6. Cool the plate to RT before reading absorbance on a plate reader at 562nm.

7. Use software or the standard curve to calculate concentrations.

A.3.2.7 Protein verification

Western Blot

1. Run SDS-PAGE with precast gel as described in protocol for SDS-PAGE (A.3.2.5), with sample sizes of 10 ng, 20 ng, 50 ng and 100 ng as well as a molecular weight standard.

2. Use the Trans-Blot® Turbo™ Transfer Pack kit, along with the Trans-Blot® Turbo™ Transfer System to blot the gel.

3. Place the membrane in the cassette followed by the gel and finally the “sponge” on top. Use a roller to make sure there are no air bubbles between gel and membrane.

4. Blot the gel for 3 min in the Trans-Blot® Turbo™ Transfer System.

5. Remove the gel, and place the membrane in a suitable container.

6. Block the membrane with 5% (0.1% Twin 20) for 1 h at 4°C.

7. Incubate membrane with primary rabbit anti-SMUG AB 1:2000 in 5% PBST milk at 4°C overnight.

8. Wash the blot with 5% PBST (0.1% Tween 20) 3 times × 10min on an orbital shaker.

9. Incubate blot with secondary Goat anti rabbit-IgG-HRP ab 1:2000 in 5% PBST milk for 1,5 h at RT.

10. Wash blot 3 times × 10min again.

11. Use the SuperSignal™ West Pico Kit to develop the blot.

12. Ananlyse the blot on ChemiDoc.