SpeedyGenes: exploiting an improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution

Andrew Currin[1,2,3], Neil Swainston[1,3,4] Philip J. Day[1,3,5,] Douglas B. Kell[1,2,3]

[1]Manchester Institute of Biotechnology, The University of Manchester, 131, Princess St, Manchester M1 7DN, United Kingdom
[2]School of Chemistry, The University of Manchester, Manchester M13 9PL, United Kingdom
[3]Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM), The University of Manchester, 131, Princess St, Manchester M1 7DN, United Kingdom
[4]School of Computer Science, The University of Manchester, Manchester M13 9PL, United Kingdom
[5]Faculty of Medical and Human Sciences, The University of Manchester, Manchester M13 9PT, United Kingdom

Abstract

Gene synthesis is a fundamental technology underpinning much research in the life sciences. In particular, synthetic biology and biotechnology utilise gene synthesis to assemble any desired DNA sequence, which can then be incorporated into novel parts and pathways. Here, we describe SpeedyGenes, a gene synthesis method that can assemble DNA sequences with greater fidelity (fewer errors) than existing methods, but that can also be used to encode extensive, statistically designed sequence variation at any position in the sequence to create diverse (but accurate) variant libraries. We summarise the integrated use of GeneGenie to design DNA and oligonucleotide sequences, followed by the procedure for assembling these accurately and efficiently using SpeedyGenes.

Key words

directed evolution – error correction – gene synthesis – protein libraries – synthetic biology

Introduction

Gene synthesis is an important driving force behind the developing disciplines of synthetic biology and biotechnology. Reducing cost and increased throughput has enabled the emergence of synthetic biology, where any desired DNA sequence can be synthesised and then assembled into pathways and genomes. However, this technology is hindered by the frequency at which errors occur in the synthesised sequence, an issue that generally arises from the oligonucleotide building blocks it is assembled from. Consequently, there is a strong demand for improvements in the methodology to further increase the accuracy and efficiency of this process. In addition, an important requirement of biotechnology and synthetic biology is the ability to encode mutations to create variant libraries for screening of novel or altered functions. In this chapter we outline SpeedyGenes [1], an improved gene synthesis method that can assemble DNA with a greater accuracy (fewer errors) than existing methods, but can also be extended generate stable variant libraries *de novo* using mixed base codons.

Most enzymes have a rather modest catalytic activity [2], primarily since natural evolution often did not have the need to select for significantly greater activity within its biological context [3,4]. However, many biotechnology applications require the heterologous expression of proteins and pathways [5], and these must undergo further selection to optimise the system for its new objective(s) [6]. This has led to the development of 'directed protein evolution' (e.g. [7-9,4]), an iterative process of mutagenesis and selection [10] (Fig 1), which selects for one or more fitnesses (fitness functions) that are of interest not to the organism but to the human experimenter. Such fitness or objective functions for enzymes often include substrate specificity, stereoselectivity, k_{cat} and thermostability [4], albeit the 'first rule' of directed evolution has been said to be that 'you get what you select for (even if you did not mean to)'.

Classically, protein diversity could be generated at the DNA level using random mutagenesis methods, primarily error-prone PCR (epPCR, [11]) or 'DNA shuffling' (the equivalent of recombination [12]). However, a fundamental problem is the simply vast extent of the search space of possible proteins [13,14,10], 20^{300} or $\sim 10^{390}$ for an approximately average sized protein of 300 amino acids. Purely random changes in sequence simply effect a local search, and substantial random mutations (that might widen the search) lead to the production of stop codons (3 per 64 residues) and premature truncation [15]. Additionally, epPCR is also limited by considerable bias towards transition mutations and also these single-base mutations cannot encode all possible amino acids [16].

However, the advent of large-scale synthetic DNA methods, even at the whole genome level [17]), means that it is now possible to target specific residues with specific mutations, which thereby circumvents many of the above problems and permits a massively improved means to navigate these very large search spaces [4]. We have developed both computational (GeneGenie [18]) and experimental (Speedygenes [1]) strategies for performing this. Our novel experimental workflow simultaneously corrects the sequence errors in the full sequence while

allowing the incorporation of any desired variation at precise residues (and at potentially enormous rates). The results of a given generation can then be used to design (again statistically) the kinds of sequence variation one might desire for subsequent generations (as we have done for aptamers [19]). It is this 'eyes open' strategy that permits the intelligent navigation of (even highly epistatic) protein search spaces [4].

The purpose of this Chapter is thus to provide a 'hands-on' guide to the use of GeneGenie and SpeedyGenes, both for efficient gene synthesis and in efficient experimental Directed Evolution studies.

Materials

- High-fidelity DNA polymerase, preferably hot start.
- dNTPs 2mM mix
- DNase-free high purity water
- DNA oligonucleotides
- Surveyor endonuclease kit (Integrated DNA Technologies)
- Thermal cycler
- DNA gel or capillary electrophoresis equipment
- Nanodrop (Thermo Scientific) or other spectrophotometer
- Set of pipettes
- PCR tubes
- Microcentrifuge tubes
- DNA purification kits, including PCR cleanup and gel extraction kits

Methods

3.1 Design of DNA oligonucleotides

The online tool GeneGenie (http://www.gene-genie.org) [18] has been designed to facilitate the design of oligonucleotides for assembly using the SpeedyGenes method. At the outset of the project it is advised to plan the full experimental cloning approach, including the target plasmid and cloning strategy. Once these have been defined, the 5' and 3' sequences required for downstream cloning can be inputted into GeneGenie to be included in the optimised oligomer design.

1. Input the desired amino acid sequence into GeneGenie, together with the 5' and 3' DNA sequences required for downstream applications (Fig 4, see Notes 1 and 2).
2. Set the design parameters, including desired melting temperature for annealing and target organism for expression (see Note 3).
3. Mixed base codons can also be inputted at this stage in the design of variant libraries for directed evolution (Fig 5, see Note 4).

3.2 Assembly of non-variant sequences

3.2.1 Intermediate block synthesis

For gene sequences of <600bp it is possible to assemble these in one PCR, however for sequences >600bp it is required to perform synthesis of multiple intermediate blocks in parallel (Fig. 2, see Notes 5 and 6).

1. Reconstitute the oligonucleotides to 100μM in high purity DNase-free water (see Note 7).
2. Create a 10 μM stock for each of the oligonucleotides to be used as PCR primers for each of the intermediate blocks (e.g. for a block containing 6 oligonucleotides, numbers 1 and 6 are the primers).
3. To create the template, pool the remaining oligonucleotides (e.g. numbers 2 to 5) into a 600 nM mixture. This is a 1/166 dilution for each oligonucleotide used (from the 100 μM stock).
4. Set up the PCR as follows (see Notes 8 and 9):

	Volume
Template (600 nM mix)	2.5 μL
Forward primer (10 μM)	3 μL
Reverse primer (10 μM)	3 μL
dNTP (2 mM)	5 μL
DNA polymerase buffer	As specified
High-fidelity DNA polymerase (hot start)	As specified
DNase-free water	Up to 50 μL

5. Carry out the PCR using a thermal cycler using the following settings:

	Temperature	**Time**

Initial denaturation	98°C	120 s
35 cycles	98°C	10 s
	60°C (or other T_m)	20 s
	72°C	20 s
Hold	4-10°C	

6. Purify the PCR products using a DNA purification kit following the manufacturer's instructions. Elute the DNA in 20 µL elution buffer.
7. Analyse the purified products using electrophoresis and measure the concentration using a spectrophotometer (see Notes 10 and 11).

3.2.2 Endonuclease digestion for error correction

1. Dilute each purified intermediate block to 100 ng/µL in 1x DNA polymerase reaction buffer (see Note 12), in a total volume of 10 µL.
2. Perform the denaturation and slow hybridisation to create mismatches:

	Temperature	**Ramp Rate**	**Time**
Initial denaturation	95°C		2 min
Slow hybridisation	85°C	2°C/s	1 min
	Lowered in 10°C intervals down to 25°C	0.3°C/s	1 min at every 10°C interval
Hold	4°C		

3. Create the endonuclease enzyme mix (see Note 13).

	Volume
Surveyor endonuclease	2 µL
Surveyor enhancer	1 µL
1x DNA polymerase buffer	As specified
DNase-free water	Up to 5 µL

4. Combine 5 µL of the hybridised DNA with 5 µL of enzyme mix and gently mix.
5. Incubate at 42°C on a thermocycler or heat block for 2 h.
6. Purify the digest products using a DNA purification kit, eluting with 10-20 µL elution buffer (see Note 14).

3.2.3 Assembly of the full-length sequence

1. To create the PCR template, add 1 µL of each of the purified block digests together.
2. For PCR primers, use the two outermost oligonucleotides in the sequence.
3. Set up the PCR reaction:

	Volume
Template	2 µL
Forward primer (10 µM)	3 µL
Reverse primer (10 µM)	3 µL
dNTP (2 mM)	5 µL
DNA polymerase buffer	As specified
High-fidelity DNA polymerase (hot start)	As specified
DNase-free water	Up to 50 µL

4. Run the PCR (see Note 15):

	Temperature	**Time**
Initial denaturation	98°C	120 s
35 cycles	98°C	10 s
	60°C (or other T_m)	20 s
	72°C	40s (or longer as required)
Hold	4-10°C	

5. Analyse the PCR products using electrophoresis (see Note 11).

3.3 Assembly of variant libraries

The three-step procedure of 1) block synthesis 2) endonuclease digestion 3) full-length assembly, is the same for the synthesis of variant and non-variant sequences. However, when oligonucleotides containing mixed base codons (here termed variant oligonucleotides) are to be used to create variant libraries, the block synthesis and full-length PCR steps are modified (Fig. 3). As such, the above protocol should be carried out using the following alternative steps.

(3.2.1)

3. To create the template for the block synthesis, pool the desired oligonucleotides (including the variant oligonucleotides) into the 600 nM mixture.
4. Carry out the PCR as in 3.2.1.4.

(3.2.3)

3. Set up the PCR including the 'spiked-in' variant oligonucleotides:

	Volume
Template	2 µL
Forward primer (10 µM)	3 µL
Reverse primer (10 µM)	3 µL
Variant oligonucleotide(s) (0.1 µM)	3 µL
dNTP (2 mM)	5 µL
DNA polymerase buffer	As specified
High-fidelity DNA polymerase (hot start)	As specified
DNase-free water	Up to 50 µL

4. Carry out the PCR as in 3.2.3.4.

3.4 Downstream cloning and *E. coli* transformation

The SpeedyGenes protocol has been developed using the In-Fusion cloning system (Clontech) to ligate into a linearised expression vector, however any cloning procedure can be used after the above procedures. Here, we outline our current procedure as an example workflow.

1. Run the SpeedyGenes PCR products on an agarose electrophoresis gel and purify the full-length product (gene) using a gel extraction purification kit.
2. Ligate the gene into a linearised plasmid using the In-Fusion cloning kit, following the manufacturer's protocol.
3. Transform the ligation product into high efficiency *E. coli* competent cells, purify the plasmid and verify the correct sequences using DNA sequencing. For screening purposes, synthesised genes can be assayed directly from the transformation culture without the need for any sequence verification.

Notes

1. It is strongly advised to input the additional 5' and 3' sequences required for cloning into the initial GeneGenie designs. This will enable GeneGenie to include these sequences in its calculations to maximise the efficiency of the PCR assembly. If these sequences are not included at this design stage then it is possible that they may anneal in an incorrect position in the sequence, thus reducing the efficiency of the assembly. Incorporating 5' and 3' sequences optimises the design of oligomers encoding *both* the cloning sequences and the gene simultaneously.
2. Check that the 5' and 3' sequences include the required sequences for any downstream procedures. These commonly include i) adding a start codon; ii) adding a stop codon; and iii) correction of the reading frame if using restriction endonucleases.
3. The length of oligonucleotides is often defined by the preferred supplier and synthesis scale required, whilst the melting temperature (T_m) is determined by the annealing temperature at which the PCR assembly will be performed. If the methodology does not yield good results, redesigning the gene using a higher T_m could resolve the issue.
4. If the design is to include mixed base codons for variant libraries, these should best be specified in the initial designs. When these are specified, GeneGenie will endeavour to design the oligonucleotides such that their positions do not fall within the overlap regions of the construct (enabling them to be mutated using a single oligonucleotide).
5. Intermediate blocks <600bp can be reliably assembled using oligonucleotides. Synthesising sequences >600bp, or containing >12 oligonucleotides, often do not assemble with high efficiency. For example, for the 747bp sequence for EGFP, this was separated into two blocks for efficient assembly.
6. For each intermediate block, the oligonucleotides at the outer 5' and 3' ends are used as PCR primers. It is therefore important that each block contains an equal number of oligonucleotides. For example, in a block containing 6 oligonucleotides, oligonucleotide 1 is the forward primer and 6 is the reverse primer.
7. This protocol has been developed using DNA oligonucleotides with a standard desalting purification. Other purifications, like HPLC, can be used but are not required. However, it is recommended that the oligonucleotides are of high quality with mass spectrometry analysis for quality control.
8. Use of a high-fidelity DNA polymerase is required for the protocol. During the final OE-PCR step (where the products from the endonuclease digest are assembled into the full-length sequence) the strong proofreading ability of this polymerase is crucial for the removal of the erroneous sequence for successful error correction. In addition, a high-fidelity polymerase will also minimise the introduction of new erroneous sequences during the PCR cycling steps. We routinely use the Q5 hot start high-fidelity polymerase (New England Biolabs).
9. Use of a hot start DNA polymerase is also recommended to prevent undesired polymerisation of the oligonucleotide template prior to PCR assembly.

If using a non-hot-start polymerase, then the reactions should be set up on ice and transferred directly to a preheated PCR machine at the start of the protocol.

10. When designing the thermal cycling protocols it is important to check the manufacturers instructions for the DNA polymerase. Importantly, this will recommend particular reaction conditions or cycling temperatures that should be used.

11. The thermal cycling conditions have a significant impact over the quality of the PCR products created. Hence, if non-specific bands are detected then optimisation of annealing temperature, annealing time and elongation time should be attempted.

12. A concentration of 70-100ng/μL is desirable for each block for the endonuclease digestion. Lower concentrations can be used successfully, but it is desirable to use roughly equal concentrations of fragments in the final PCR assembly.

13. Due to the viscosity of the enzyme mix, it is recommended to create a larger stock mixture sufficient for all the digest reactions.

14. It is preferable to elute in a low volume at this step to maximise the DNA concentration.

15. Depending on the size of the sequence to be assembled it may be necessary to increase the extension time. Refer to the manufacturers instructions for the speed of the DNA polymerase.

Concluding remarks.

The basis for the directed evolution of proteins lies in the ability to create and assay variant sequences for improved properties. Classical methods were more or less entirely empirical (random), but synthetic biology facilitates the means to create (statistically) precisely variant DNA sequences at each residue. This permits the exploitation of (theoretically desirable [20]) high mutation rates without the potential creation of stop codons. However, DNA synthesis is itself prone to errors, but standard error correction procedures might simply remove the diversity that one is seeking to create! Speedygenes [1,4] gets round this by separating the error-correction and diversity-introducing steps, with optimisation controlled by the GeneGenie [18] software. In contrast to methods such as ProSAR [21], this permits the exploration and exploitation even of epistatic interactions, and hence the intelligent navigation [4] of the very large search spaces involved.

Acknowledgements

We thank the Biotechnology and Biological Sciences Research Council for financial support (grant BB/M017702/1), Prof Nick Turner, Dr Ian Rowles and Dr Timothy Eyes for useful discussions, and Mrs Hannah Currin for preparation of figures. This is a contribution from the Manchester Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM).

Figure legends

Fig. 1. The overall procedure for an evolutionary system. At the start, an individual protein is selected and assessed for its fitness (i.e. its activity). Genetic diversity then creates a library of variants, from which improved individuals are selected. This cycle is repeated until the individual fulfils the desired fitness criteria.

Fig. 2. The SpeedyGenes workflow for non-variant sequences. Genes are first assembled from DNA oligonucleotides into intermediate blocks (an example of two blocks is shown) using overlap-extension PCR (OE-PCR). These blocks then undergo incubation with the mismatch endonuclease Surveyor to cleave sequence errors. These digested products are then pooled together and used as the template for a second OE-PCR, which assembles the full-length, error-corrected sequence.

Fig. 3. The SpeedyGenes method for synthesis of variant libraries. Variant oligonucleotides are used for the intermediate block synthesis. These mixed base sequences are heavily digested by the Surveyor nuclease. Consequently, the same variant oligonucleotides are 'spiked' into the final OE-PCR step to assemble the full-length variant sequence.

Fig. 4. The GeneGenie query page, allowing specification of the target protein sequence, 5' and 3' cloning sequences, maximum oligomer length, target melting temperature, and host organism for expression.

Fig. 5. The GeneGenie query page, showing specification of the variant codon NTN at a specific position, introducing variability at this site and ultimately generating variant libraries.

Fig. 6. A GeneGenie result page, showing the optimised oligomer design in terms of its alignment and list of oligomers to synthesise or purchase. 5' and 3' cloning sequences are underlined in the alignment, with variant codons highlighted in green. Other statistics, including those relating to optimised codon usage, are also displayed.

Fig. 7. Synthesis of a monoamine oxidase (MAO-N) and direct functional assay for catalytic activity. DNA fragments are analysed using capillary electrophoresis. (A) The 1518 bp sequence was assembled using four intermediate blocks (labelled 1-4), followed by (B) mismatch endonuclease digestion of the blocks. (C) Pooling of these digest products and assembly into the full MAO-N sequence by OE-PCR. (D) Direct ligation and expression of this gene in *E. coli* showed 76% clones with full catalytic activity (and correct DNA sequence).

Fig. 8. Synthesis of controlled variant libraries of enhanced green fluorescent protein (EGFP). Residues 66 and 145 were mutated in a reduced library using mixed base codons, mutating these codons to encode Y/H and Y/F, respectively. The variant oligonucleotides encoding these mutations (6 and 12) were (A) assembled into the intermediate blocks then (B) subject to endonuclease

digestion. (C) Variant oligonucleotides were then ‘spiked in’ for the final OE-PCR step. (D) Expression of this library in *E. coli* identified both green and blue variants when analysed under UV light.

Table 1. The IUPAC code [22] for mixed DNA bases.

Symbol	Nucleotide base
G	G
A	A
T	T
C	C
R	G, A
Y	T, C
M	A, C
K	G, T
S	G, C
W	A, T
H	A, C, T
B	G, T, C
V	G, C, A
D	G, A, T
N	G, A, T, C

Table 2. An example of the mixed base codons that can be used for generating controlled genetic diversity. The number of possible codons, inclusion of stop codons and the amino acids they encode are also highlighted.

Degenerate codon	Mixed base sequence	Encoded codons	Stop codons	Encoded amino acids	Properties
NNN	(A,T,G,C) (A,T,G,C) (A,T,G,C)	64	TAA, TAG, TGA	All	Fully randomised codon
NNK	(A,T,G,C) (A,T,G,C) (G,T)	32		All	All 20 amino acids
NDT	(A,T,G,C) (A,T,G) T	12	No	Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly	Mixture of polar, nonpolar, positive and negative charge (Reetz 2008)
NTN	(A,T,G,C) T (A,T,G,C)	16	No	Met, Phe, Leu, Ile, Val	Nonpolar residues

References

1. Currin A, Swainston N, Day PJ, Kell DB (2014) SpeedyGenes: a novel approach for the efficient production of error-corrected, synthetic gene libraries. Protein Eng Design Sel 27:273-280. doi:DOI: 10.1093/protein/gzu029
2. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50 (21):4402-4410
3. Kacser H, Burns JA (1981) The molecular basis of dominance. Genetics 97:639-666
4. Currin A, Swainston N, Day PJ, Kell DB (2015) Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Chem Soc Rev 44 (5):1172-1239. doi:10.1039/c1034cs00351a
5. Kell DB, Westerhoff HV (1986) Metabolic control theory: its role in microbiology and biotechnology. FEMS Microbiol Rev 39:305-320
6. Mendes P, Kell DB (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. Bioinformatics 14:869-883
7. Arnold FH, Volkov AA (1999) Directed evolution of biocatalysts. Curr Op Chem Biol 3 (1):54-59
8. Voigt CA, Kauffman S, Wang ZG (2001) Rational evolutionary design: The theory of *in vitro* protein evolution. In: Arnold FM (ed) Advances in Protein Chemistry, Vol 55, vol 55. Advances in Protein Chemistry. pp 79-160
9. Turner NJ (2009) Directed evolution drives the next generation of biocatalysts. Nat Chem Biol 5 (8):567-573
10. Kell DB, Lurie-Luke E (2015) The virtue of innovation: innovation through the lenses of biological evolution. J R Soc Interface 12 (2):20141183. doi:10.1098/rsif.2014.1183
11. McCullum EO, Williams BA, Zhang J, Chaput JC (2010) Random mutagenesis by error-prone PCR. Methods Mol Biol 634:103-109. doi:10.1007/978-1-60761-652-8_7
12. Stemmer WPC (1994) Rapid evolution of a protein *in vivo* by DNA shuffling. Nature 370:389-391
13. Reetz MT, Kahakeaw D, Lohmer R (2008) Addressing the numbers problem in directed evolution. Chembiochem 9 (11):1797-1804
14. Kell DB (2012) Scientific discovery as a combinatorial optimisation problem: how best to navigate the landscape of possible experiments? Bioessays 34 (3):236-244
15. Pritchard L, Corne DW, Kell DB, Rowland JJ, Winson MK (2004) A general model of error-prone PCR. J Theoret Biol 234 (4):497-509
16. Zhao J, Kardashliev T, Joelle Ruff A, Bocola M, Schwaneberg U (2014) Lessons from diversity of directed evolution experiments by an analysis of 3,000 mutations. Biotechnol Bioeng 111:2380-2389. doi:10.1002/bit.25302
17. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA, 3rd, Smith HO (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science 319 (5867):1215-1220

18. Swainston N, Currin A, Day PJ, Kell DB (2014) GeneGenie: optimised oligomer design for directed evolution. Nucleic Acids Res 12:W395-W400. doi:10.1093/nar/gku336
19. Knight CG, Platt M, Rowe W, Wedge DC, Khan F, Day P, McShea A, Knowles J, Kell DB (2009) Array-based evolution of DNA aptamers allows modelling of an explicit sequence-fitness landscape. Nucleic Acids Res 37 (1):e6
20. Oates MJ, Corne DW, Kell DB (2003) The bimodal feature at large population sizes and high selection pressure: implications for directed evolution. In: Tan KC, Lim MH, Yao X, Wang L (eds) Recent advances in simulated evolution and learning. World Scientific, Singapore, pp 215-240
21. Fox RJ, Davis SC, Mundorff EC, Newman LM, Gavrilovic V, Ma SK, Chung LM, Ching C, Tam S, Muley S, Grate J, Gruber J, Whitman JC, Sheldon RA, Huisman GW (2007) Improving catalytic function by ProSAR-driven enzyme evolution. Nat Biotechnol 25 (3):338-344
22. Nomenclature Committee of the International Union of Biochemistry (NC-IUB) (1985) Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. Eur J Biochem 150:1-5

Directed Evolution overview

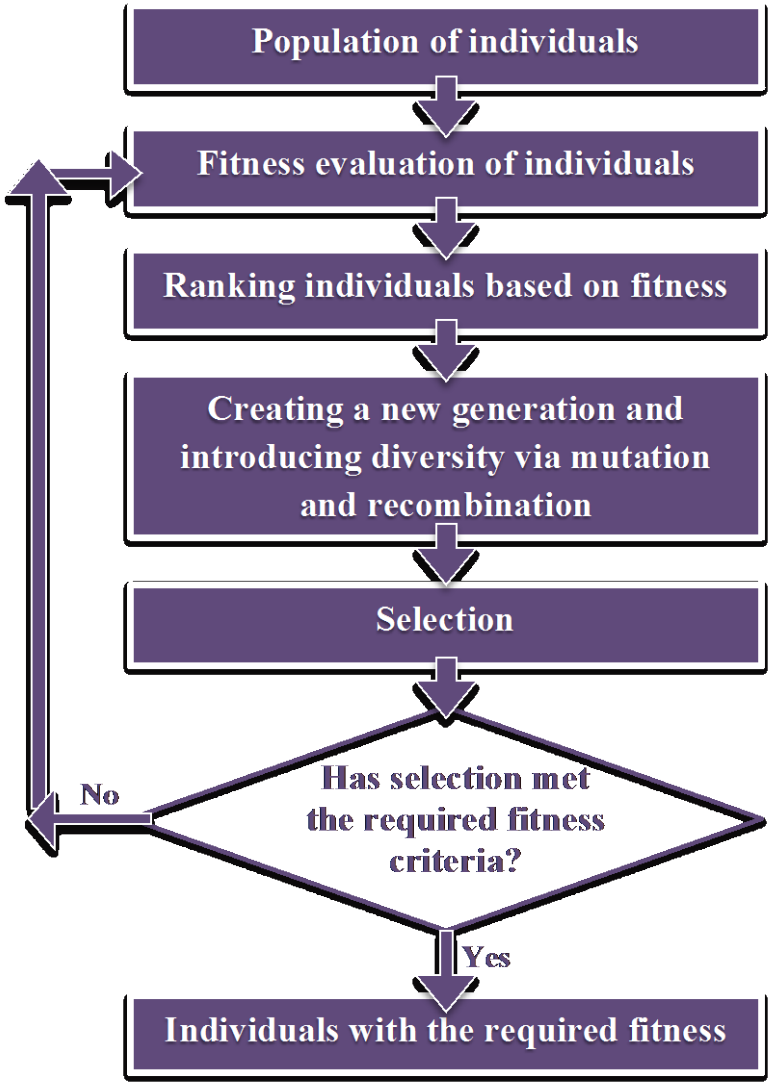

Fig 1

SpeedyGenes overview

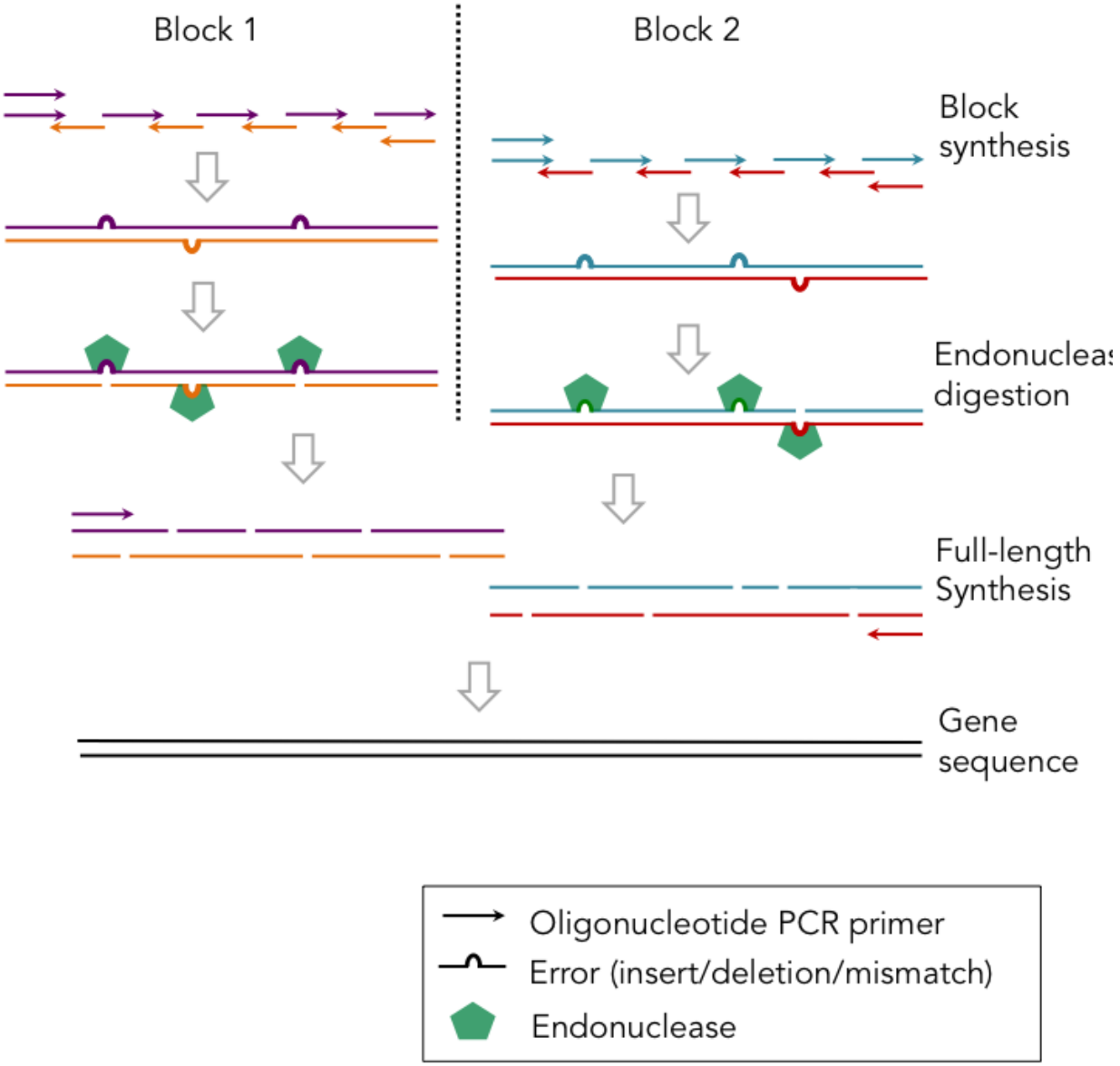

Fig 2

SpeedyGenes for variants

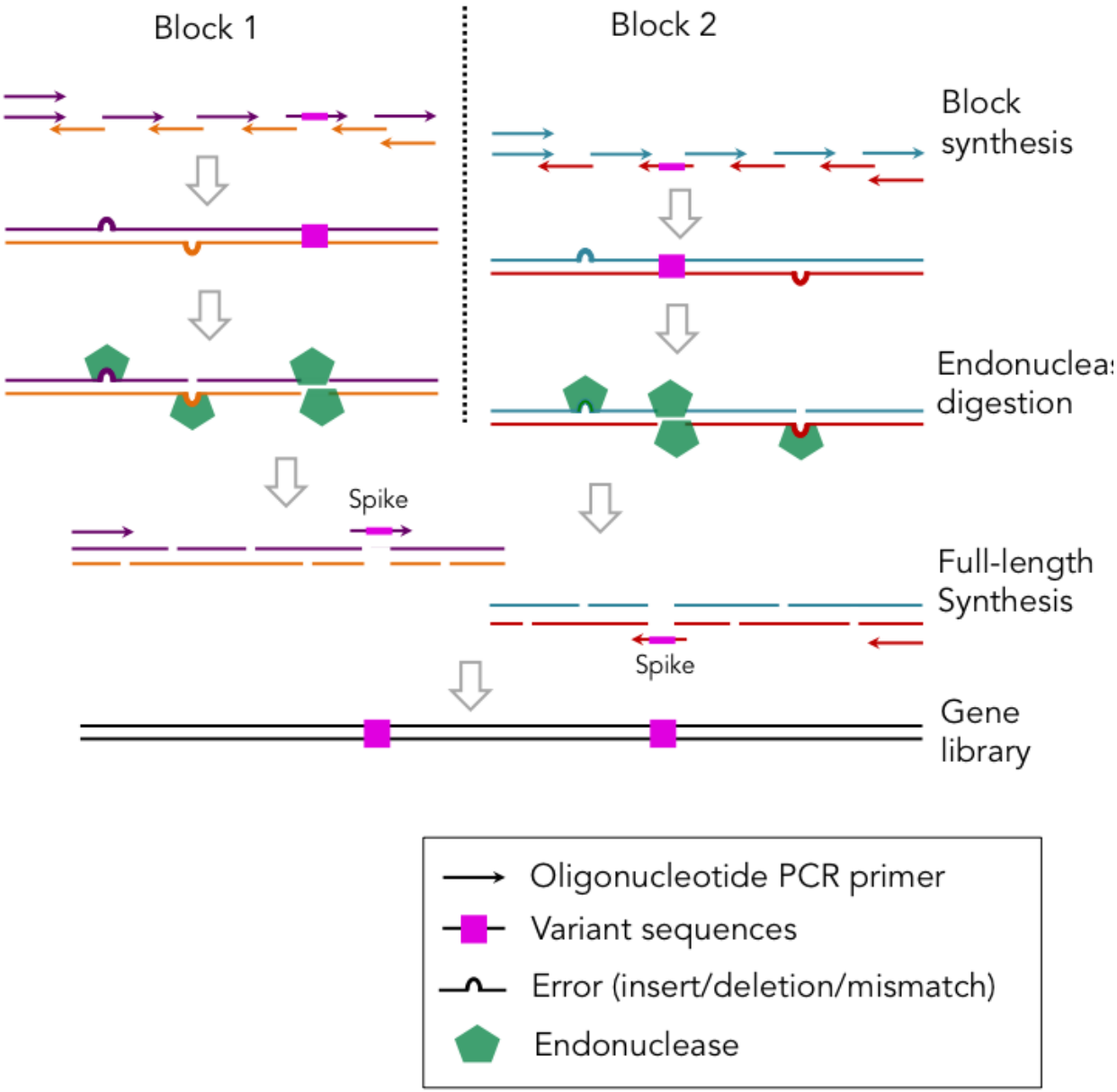

Fig 3

GeneGenie input page

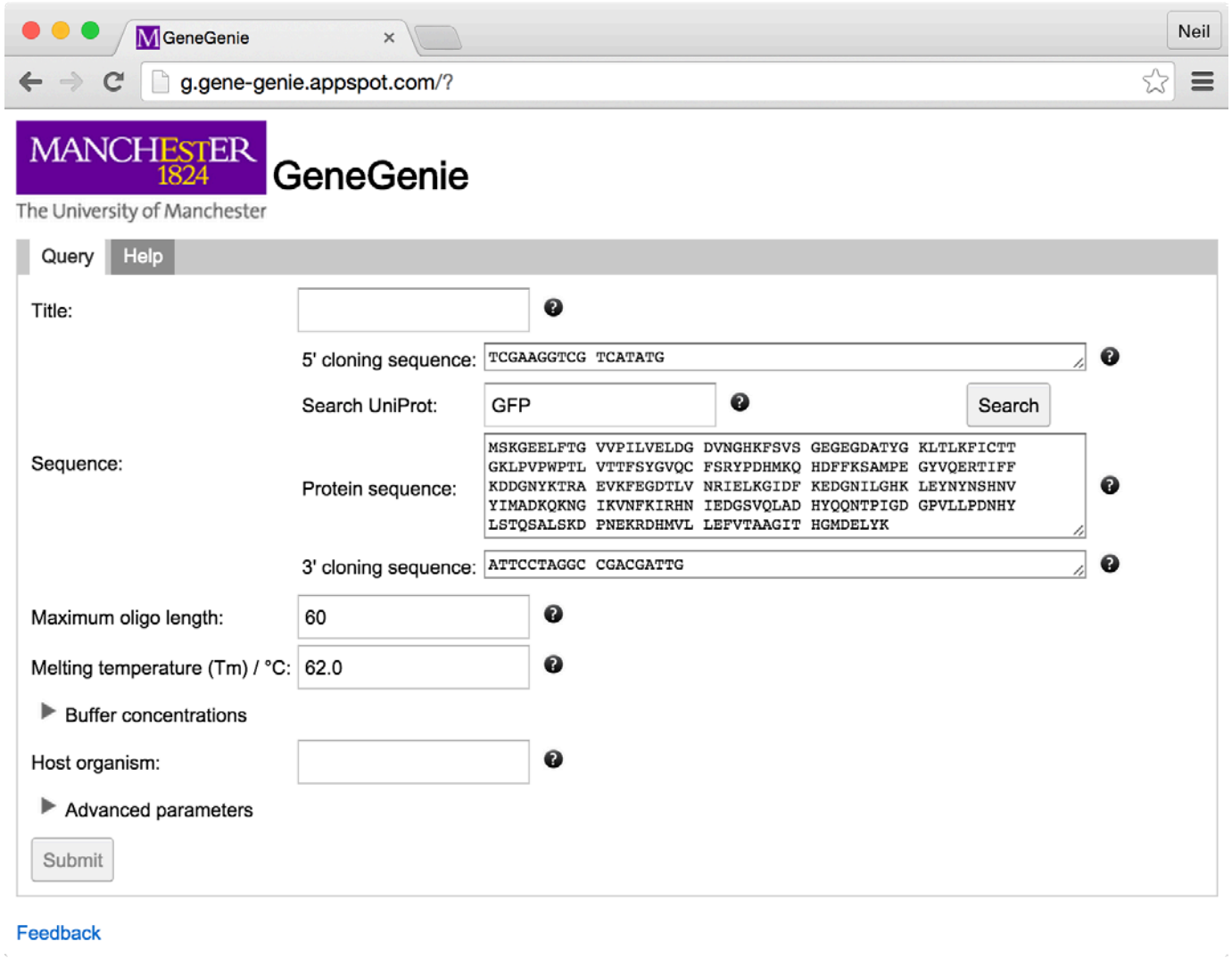

Fig 4

Inputting GeneGenie variants

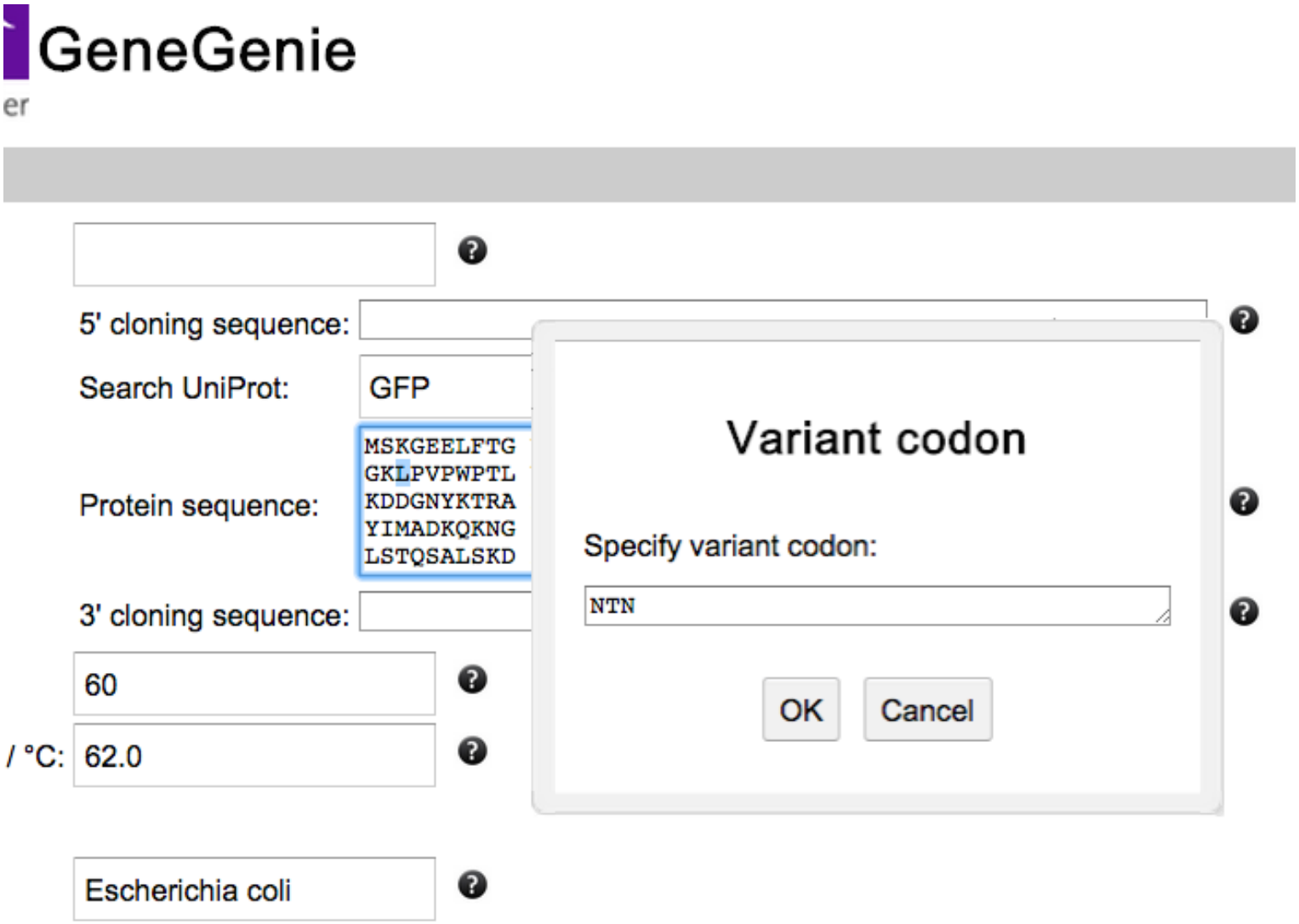

Fig 5

A typical GeneGenie output

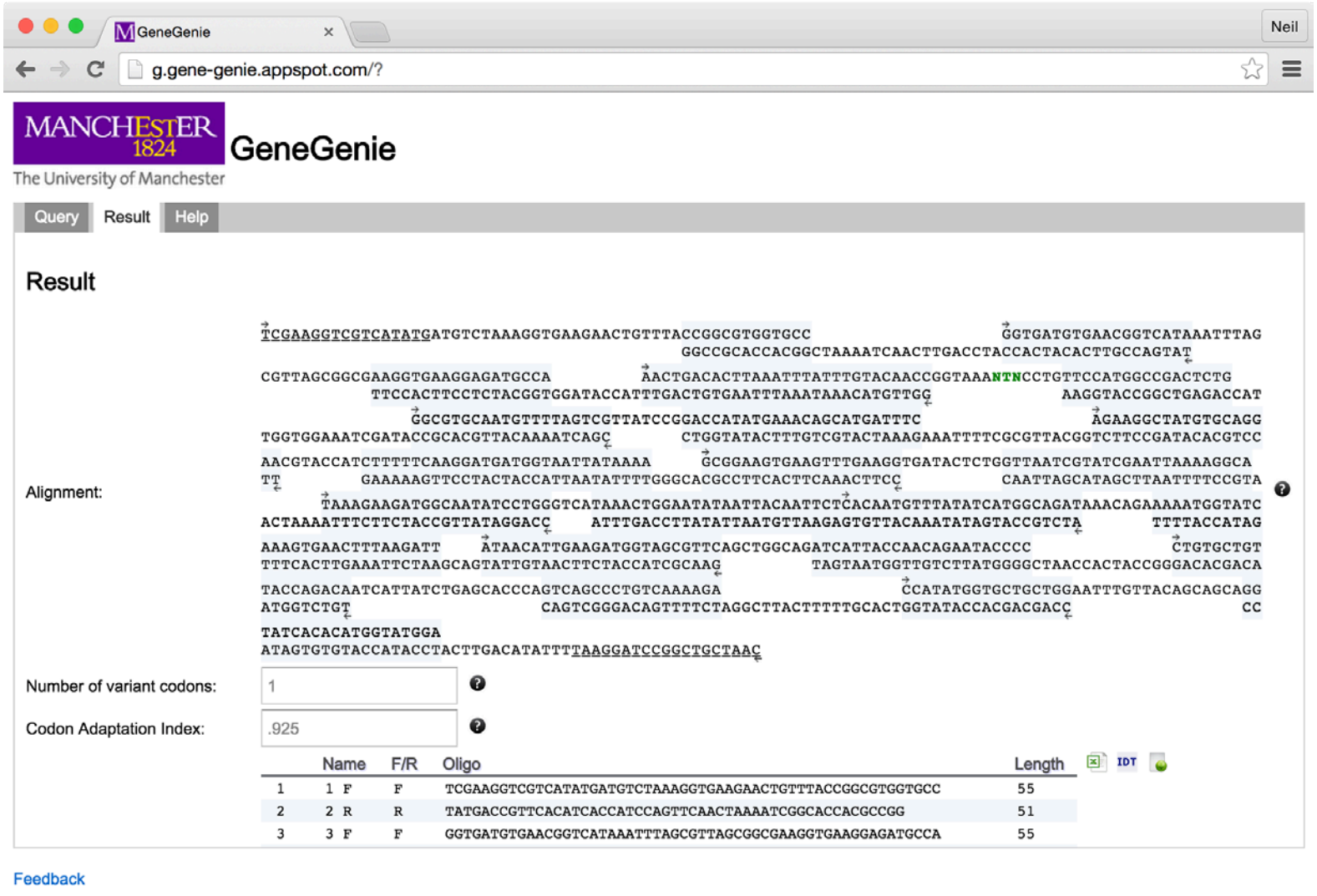

Fig 6

SpeedyGenes synthesis of MAO-N

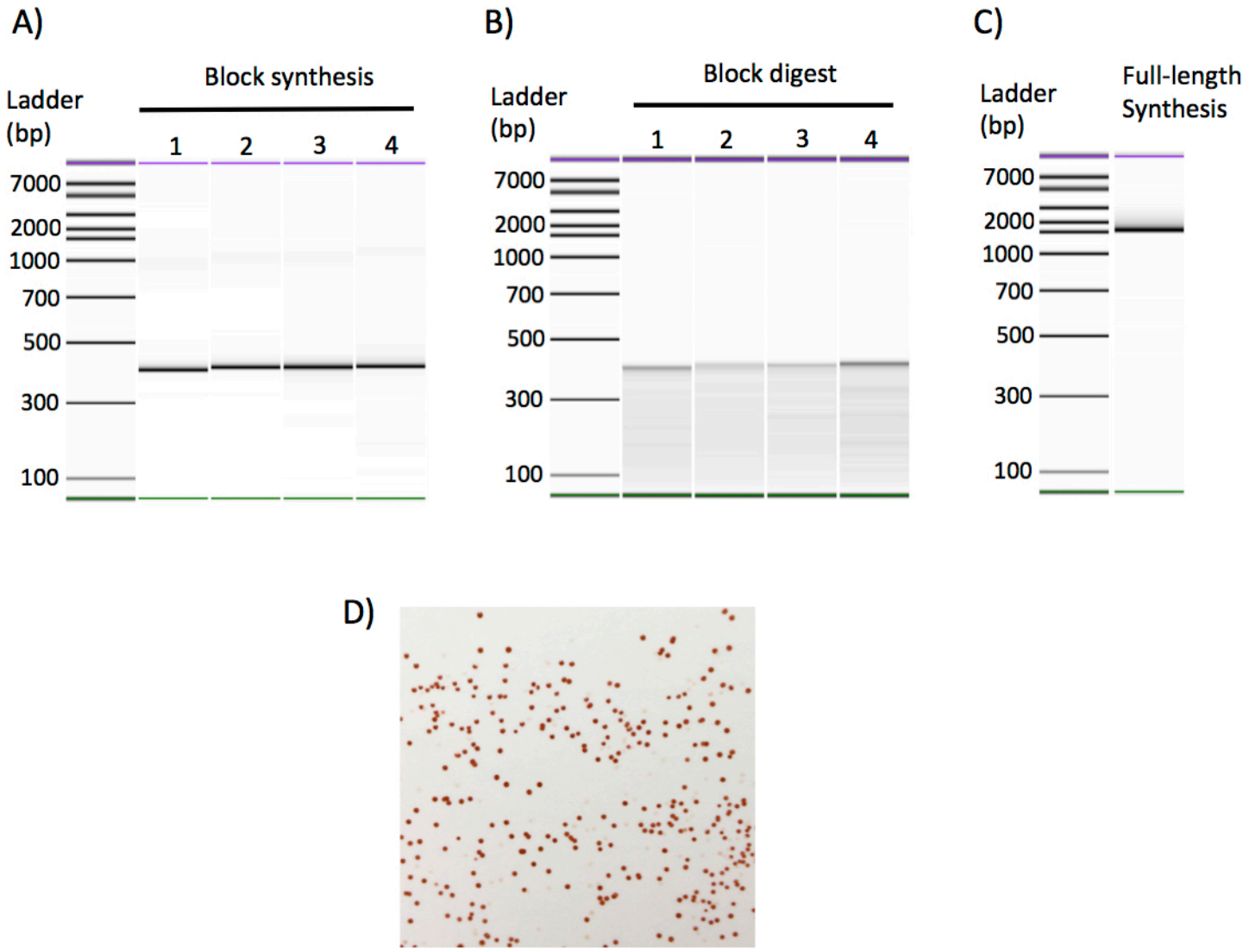

Fig 7

SpeedyGenes variants – EGFP/BFP library

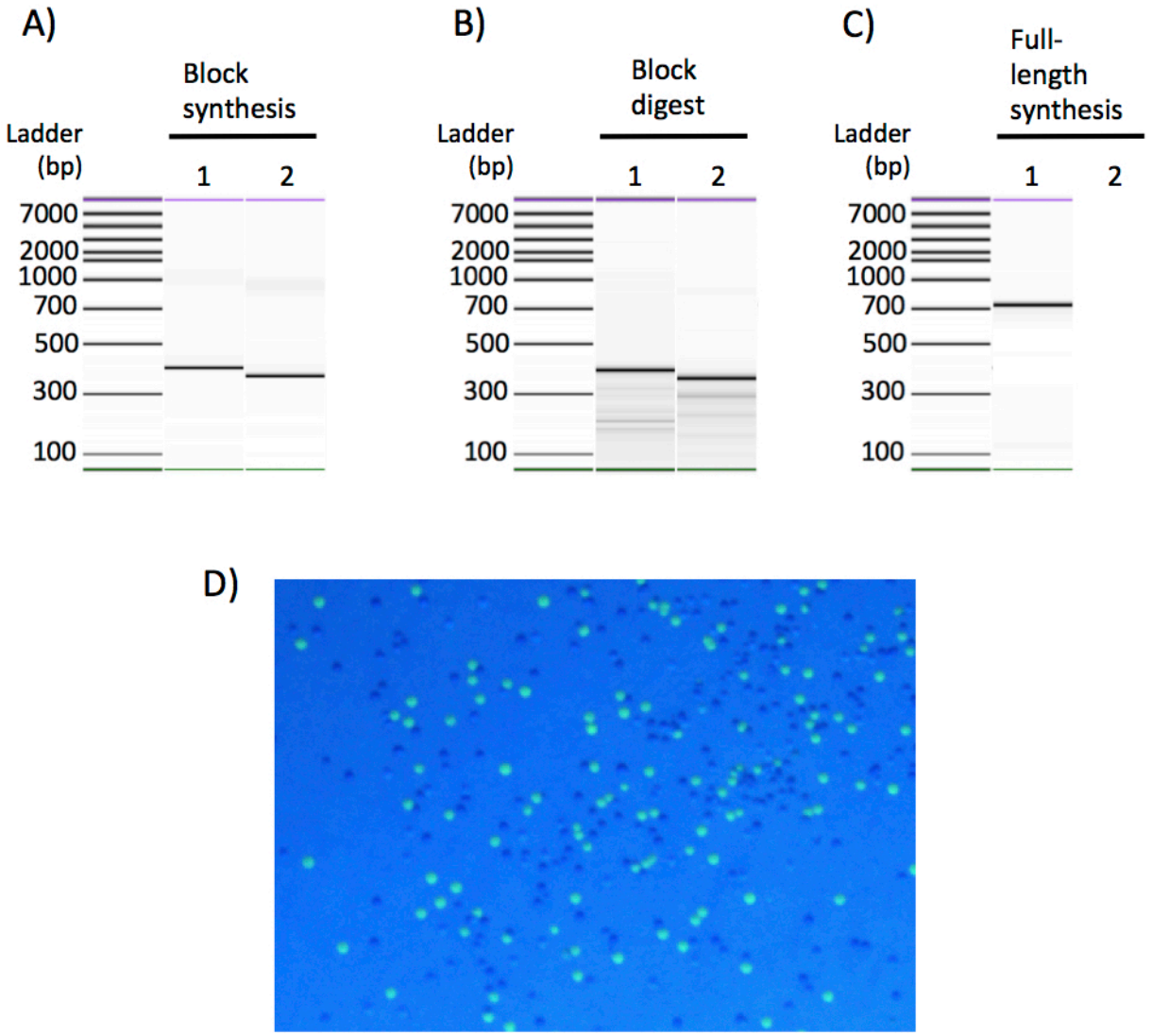

Fig 8