Claims

1. A method of synthesizing a DNA molecule having a desired sequence comprising:

a) annealing at least two oligonucleotides to an anchor strand so that the at least two oligonucleotides annealed to the anchor strand abut one another on the anchor strand;

wherein the at least two oligonucleotides each comprise a primer binding site on a 3' or 5' end, and a variable sequence on the opposing 5' or 3' end, and a conserved flanking sequence in between the primer binding site and the variable sequence; and

wherein the anchor strand comprises conserved flanking sequences complementary to those on the at least two oligonucleotides, and further comprises at least one variable sequence, wherein at least a portion of the at least one variable sequence on the anchor strand is complementary to at least a portion of the variable sequences on the at least two oligonucleotides;

b) ligating the at least two oligonucleotides annealed to the anchor strand to produce a first dsDNA molecule;

c) performing an amplification step on the first dsDNA molecule having a desired sequence and comprising a conserved flanking sequence inside each of the 3' and 5' ends, and a variable sequence inside the conserved flanking sequences.

2. The method of claim 1 further comprising contacting the first dsDNA molecule with a restriction endonuclease to produce first dsDNA fragments comprising 3' and/or 5' overhang sequences comprising a portion of the variable sequence from the first dsDNA molecule, providing at least one first additional dsDNA fragment comprising a 3' and/or 5' overhang sequence that is at least partially complementary to an overhang sequence of at least one of the first dsDNA fragments;

annealing the first dsDNA fragments and at least one first additional dsDNA fragment by the 3' and/or 5' overhang sequences; and

ligating the annealed dsDNA fragments to produce a second dsDNA molecule comprising a conserved flanking sequence inside each of the 3' and 5' ends, and a variable sequence inside the 3' and 5' conserved flanking sequences that is longer than the variable sequence on the first dsDNA molecule.

3. The method of claim 2 wherein the at least one first additional dsDNA fragment is the product of a parallel DNA synthesis reaction, further wherein the first dsDNA molecule has a recognition site for a restriction endonuclease on the 5' or 3' side of the molecule, and the first additional dsDNA fragment is derived from restriction cleavage of a dsDNA molecule having a recognition site for a restriction endonuclease on the opposing 3' or 5' side of the molecule.

4. The method of claim 3 further comprising contacting the at least one second dsDNA molecule with a restriction endonuclease to produce a plurality of second dsDNA fragments comprising 3' and/or 5' overhang sequences and a conserved flanking sequence inside each of the 3' or 5' ends;

providing at least one second additional dsDNA fragment comprising a 3' and/or 5' overhang sequence that is at least partially complementary to an overhang sequence of at least one of the second dsDNA fragments;

annealing the plurality of second dsDNA fragments to the at least one second additional dsDNA fragment by the 3' and/or 5' overhang sequence(s); and

performing a step of ligation to produce a third dsDNA molecule comprising a conserved flanking sequence on the 3' and 5' ends, and a variable sequence inside the conserved flanking sequences that is longer than the variable sequence of the second dsDNA molecule.

5. The method of claim 4 wherein the at least one second additional dsDNA fragment is the product of a parallel DNA synthesis reaction, further wherein the second dsDNA molecule has a recognition site for a restriction endonuclease on the 5' or 3' side of the molecule, and the second additional dsDNA fragment is derived from restriction

cleavage of a dsDNA molecule having a recognition site for a restriction endonuclease on the opposing 3' or 5' side of the molecule.

6. The method of any one of claims 4-5 further comprising reacting the at least one third dsDNA molecule with a restriction endonuclease to produce a plurality of third dsDNA fragments comprising 3' and/or 5' overhang sequences and a conserved flanking sequence inside each of the 3' or 5' ends;

providing at least one third additional dsDNA fragment comprising a 3' and/or 5' overhang sequence that is at least partially complementary to an overhang sequence of at least one of the third dsDNA fragments;

annealing the plurality of third dsDNA fragments to the at least one third additional dsDNA fragment by the 3' and/or 5' overhang sequences; and

performing a step of ligation to produce a fourth dsDNA molecule comprising a conserved flanking sequence on the 3' and 5' ends, and a variable sequence inside the conserved flanking sequences that is longer than the variable sequence of the third dsDNA molecule.

7. The method of claim 6 wherein the at least one third additional dsDNA fragment is the product of a parallel DNA synthesis reaction, further wherein the third dsDNA molecule has a recognition site for a restriction endonuclease on the 5' or 3' side of the molecule, and

the first additional dsDNA fragment is derived from restriction cleavage of a dsDNA molecule having a recognition site for a restriction endonuclease on the opposing 3' or 5' side of the molecule.

8. The method of any one of claims 1-7 wherein the first dsDNA molecule comprises a variable sequence of 8-12 base pairs.

9. The method of claim 7 wherein:

the second dsDNA molecule comprises a variable sequence of 14-18 base pairs; or

the third dsDNA molecule comprises a variable sequence of 24-32 base pairs; or

the fourth dsDNA molecule comprises a variable sequence of 90-110 base pairs; or.

the at least two oligonucleotides have a variable sequence of 4-6 nucleotides; or

the amplification step is performed by the polymerase chain reaction (PCR); or

the variable sequence of the anchor strand is equal in length to the lengths of the variable sequences on the at least two oligonucleotides; or

the anchor strand comprises a variable sequence present in between the two sequences complementary to the conserved flanking sequences on the at least two oligonucleotides; or

the at least two oligonucleotides bound to the anchor strand abut one another on the anchor strand at their variable sequences; or

the portion of the variable sequence on the anchor strand that is complementary to the conserved flanking sequence on the at least two oligonucleotides comprise 2-6 nucleotides or 14-18 nucleotides or 26-30 nucleotides or 90-110 nucleotides; or

the at least two oligonucleotides and anchor strand further comprise a recognition site for a restriction endonuclease; or

the restriction endonuclease is a Type IIS endonuclease; or

the step of ligation occurs spontaneously; or

the anchor strands comprise 4-6 degenerate nucleotides; or

the degenerate nucleotides comprise a universal or randomized base; or

the DNA molecule of desired sequence has an error rate of less than 1 base pair per 2,000 versus the desired sequence; or

the DNA molecule of desired sequence is assembled from a library of fewer than 20,000 or 10,000 members; or

the primer binding sites comprise universal primer binding sites.

10. A composition comprising at least two oligonucleotides, each comprising a primer binding site on a 3' or 5' end, and a variable sequence on the opposing 5' or 3' end, and a conserved flanking sequence in between the primer binding site and the variable sequence; and

wherein an anchor strand comprising sequences complementary to the conserved flanking sequences on the at least two oligonucleotides, and further comprising at least one variable sequence, wherein at least a portion of the at least one variable sequence on the anchor strand is complementary to at least a portion of the variable sequences on the at least two oligonucleotides.

11. The composition of claim 10 wherein:

the anchor strand comprises sequences complementary to the conserved flanking sequences on the at least two oligonucleotides at its 3' and 5' ends; or

the anchor strand comprises the variable sequence in between the two sequences complementary to the conserved flanking sequence; or

the primer binding sites comprise universal primer binding sites.

12. A method of storing data in a DNA sequence comprising:

determining a sequence of DNA that encodes a non-genetic message according to a coding scheme that translates the non-genetic message from a reference language into a DNA sequence and vice versa;

synthesizing the sequence of DNA that encodes the non-genetic message according to the method of any one of claims 1-9; and

thereby store data in a DNA sequence.

13. The method of claim 12 wherein the DNA is encoded in bytes of 4 or 5 or 6 nucleotides.

14. A method of synthesizing a DNA sequence encoding a guide RNA comprising:

determining a sequence of DNA that encodes a guide RNA;

synthesizing the sequence of DNA that encodes the guide RNA according to any one of claims 1-9.

15. An oligonucleotide library comprising 1,536 distinct locations, comprising

a. 1,024 locations comprising an oligonucleotide having a unique variable sequence of non-degenerate nucleotides; and

b. an additional 512 distinct locations, each of the 512 locations comprising an anchor strand having a variable sequence comprising at least three nondegenerate nucleotides and at least four degenerate nucleotides.

16. The oligonucleotide library of claim 15 wherein the oligonucleotides comprise a primer binding site on a 3' or 5' end, and a variable sequence on the opposing 5' or 3' end, and a conserved flanking sequence in between the primer binding site and the variable sequence; and

the anchor strands comprising conserved flanking sequences complementary to those on the at least two oligonucleotides, and further comprises at least one variable sequence, wherein at least a portion of the at least one variable sequence on the anchor strand is complementary to at least a portion of the variable sequences on the at least two oligonucleotides.

17. The oligonucleotide library of claim 16 wherein the 512 distinct locations comprise anchor strand oligonucleotides having every possible sequence for the variable sequence of the anchor strand.

18. The oligonucleotide library of claim 17 wherein the variable sequence of the anchor strand comprises five or six nucleotides.

19. The oligonucleotide library of claim 15, comprising 4,608 distinct locations, comprising

a. 4,096 locations comprising an oligonucleotide having a unique variable sequence comprising non-degenerate nucleotides;

b. an additional 512 distinct locations, each of the at least 512 locations comprising an anchor strand having a variable sequence comprising at least three non-degenerate nucleotides and at least five degenerate nucleotides.

20. The oligonucleotide library of claim 19 wherein the oligonucleotides comprise a primer binding site on a 3' or 5' end, and a variable sequence on the opposing 5' or 3' end, and a conserved flanking sequence in between the primer binding site and the variable sequence; and

the anchor strands comprising conserved flanking sequences complementary to those on the at least two oligonucleotides, and further comprises at least one variable sequence, wherein at least a portion of the at least one variable sequence on the anchor strand is complementary to at least a portion of the variable sequences on the at least two oligonucleotides.

21. The oligonucleotide library of claim 20 wherein the 512 distinct locations comprise anchor strand oligonucleotides having every possible sequence for the variable sequence of the anchor strand.

22. The oligonucleotide library of claim 21 wherein the variable sequence of the anchor strand comprises five or six nucleotides.