ANSA
BIOTECHNOLOGIES

Access More Sequences with Ansa Biotechnologies

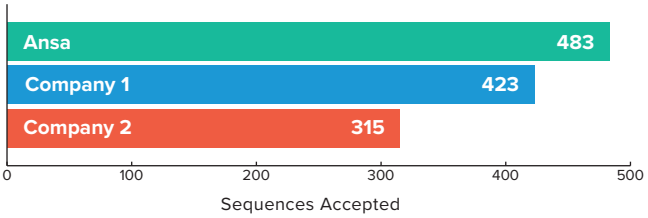
Modern biological experiments require the ability to interrogate large and diverse sets of DNA. The collection of naturally occurring DNA sequences is immense, and even larger is the set of potential synthetic DNA sequences. But DNA synthesis companies struggle to deliver on a vast portion of possible DNA design. The constraints their processes impose on DNA accessibility have reduced the effectiveness of countless experiments.

Ansa's initial products are up to 600 base pairs long, representing a large universe of potential sequences. To date, no DNA synthesis company has devised a process that can access all DNA permutations, but Ansa's enzymatic synthesis technology greatly increases the number and diversity of sequences that can now be produced. Ansa's DNA synthesis approach has the industry's lowest error rate, which allows us to directly create 600 bp oligonucleotides. These long oligonucleotides enable the generation of DNA with complex features that cannot be constructed by assembling short oligonucleotides.

Examples of Accessible Sequences

Human Genome

From a random sample of 500 base pair regions of the hg18 human genome reference, Ansa accepted 96% of the sequences for clonal gene synthesis while other DNA synthesis companies accepted just 63% to 85% of these same sequences.



Extreme GC Content

The human genome was randomly sampled to identify 600 base pair regions containing a 50 base pair window of high or low GC content. These regions were then filtered to 144 sequences representing an even distribution of local GC content from 10% to 20% and 80% to 90% across the set of sequences. Ansa successfully synthesized 96% of these sequences as gene fragments. Other DNA synthesis companies reject 25% to 58% of these sequences.

Promoters and 5'UTRs

Promoters and UTRs often contain features that can complicate manufacturing such as low GC regions, repeat regions for protein binding, and inverted repeats that can form secondary structures. A random sample of promoters with 5'UTRs (the 600 bp directly 5' of the start codon) from the human, maize, and *Arabidopsis* genomes were synthesized by Ansa with a 90% success rate at generating gene fragments. Other DNA synthesis companies reject 20% to 65% of these same sequences.

Adeno-Associated Virus (AAV) Inverted Terminal Repeats (ITRs)

AAV ITRs have a characteristic T-shaped secondary structure which consists of around 150 base pairs. Clonal gene synthesis of the ITRs from AAVs serotypes 1-13 were manufactured by Ansa and all but one were successful. Other DNA synthesis companies reject all of these sequences at the outset or fail to deliver them.

GC Content

	Accepted local GC % within any 50 base pair region
Clonal Genes	10%-94%
Gene Fragments	6%-90%

Mixing regions of high and low GC content within a sequence is allowed.

Homopolymers

Homopolymers present a unique challenge as their length is difficult to measure. Ansa can synthesize long homopolymers and we employ homopolymer-specific QC to assess them.

Homopolymer Acceptance Criteria

	Sizes of homopolymers accepted	
	Gene Fragments	Clonal Genes
A or T homopolymers	≤ 140 bp	≤ 80 bp
C or G homopolymers	≤ 20 bp	≤ 16 bp

Ansa can also manufacture multiple homopolymers of the same base in close distance. For calculating the allowed homopolymer length in a local region, homopolymers longer than 4 base pairs are considered.

	Multiple homopolymer limits	
	Gene Fragments	Clonal Genes
A or T homopolymers	≤ 140 bp in a 266 bp window	≤ 80 bp in a 152 bp window
C or G homopolymers	≤ 20 bp in a 38 bp window	≤ 16 bp in a 30 bp window

Homopolymer QC

Ansa uses next generation sequencing (NGS) to assess the quality of all the DNA we produce, including measuring the length of all homopolymers. Ansa can determine the length of short homopolymers accurately and guarantees precise sequence analysis of molecules containing this feature.

	Range of homopolymer lengths subject to exact length QC
A or T homopolymers	≤ 19 bp
C or G homopolymers	≤ 9 bp

All sequencing technologies have increased error rates in homopolymers, making it difficult to have high confidence when measuring the length of a long homopolymer. Ansa can produce long homopolymers, but due to limitations in sequencing technology we cannot guarantee long homopolymers will be exactly the length requested.

	Range of homopolymer lengths subject to variable length QC	
	Gene Fragments	Clonal Genes
A or T homopolymers	20-140 bp	20-80 bp
C or G homopolymers	10-20 bp	10-16 bp

Ansa guarantees that long homopolymers in DNA fragments are within ± 10% of the length requested and that long homopolymers in clonal DNA constructs are within ± 20% of the length requested for A or T homopolymers and within ± 10% for C or G homopolymers.

	Allowed homopolymer length variation	
	Gene Fragments	Clonal Genes
A or T homopolymers	± 10%	± 20%
C or G homopolymers	± 10%	± 10%

Repeats

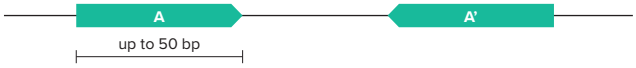
Tandem Direct Repeats

Repeated sequences that are adjacent to each other and in the same orientation are called tandem direct repeats. Tandem direct repeats, especially those with a repeat unit 2-6 base pairs long, can be difficult to manufacture. Ansa can manufacture short repeat sequences.

	Length of repeat sequence accepted
Dinucleotide repeat	≤ 20 bp
Trinucleotide repeat	≤ 30 bp
Tetranucleotide repeat	≤ 40 bp
Pentanucleotide repeat	≤ 50 bp
Hexanucleotide repeat	≤ 60 bp

Inverted Repeats

A repeated sequence where one copy is the reverse-complement of the other is an inverted repeat. There are often additional bases between the two regions of repeated sequence. Ansa accepts sequences where the length of one side of the inverted repeat is up to 50 base pairs.



Multiple inverted repeats often occur within a region of DNA. For example, in adeno-associated virus inverted terminal repeats, a "T" shaped secondary structure is the result of 3 sets of inverted repeats. Multiple inverted repeats will be accepted if one side of every individual repeat is 50 bp or less.

Conclusion

The need for accessing complex DNA sequences is imperative to enable continued advancements across healthcare, life science research, and other industries. Ansa's enzymatic synthesis technology marks a significant evolution in the ability to access complexity for fragment and clonal products.