

# De Novo Sequencing of Complex Genomes

*One system, one complete solution for de novo sequencing and high quality assembly*



Take your complex genome de novo sequencing projects to new lengths with the GS FLX Titanium Series. Combine 400-500 base shotgun and multi-span paired end reads (3 Kb, 8 Kb, 20 Kb) with advanced assembly tools to rapidly generate high quality draft assemblies.

Complex Genome Assembly			
Organism	<i>Drosophila melanogaster</i>	<i>Arabidopsis thaliana</i>	<i>Cucumber</i>
Genome Size	175 Mb	157 Mb	376 Mb
N50 Contig Size	33 Kb	37 Kb	30 Kb
N50 Scaffold Size	5.4 Mb	4.6 Mb	1.1 Mb
Largest Scaffold	10.9 Mb	9.3 Mb	5.6 Mb
Oversampling			
Shotgun	12x	17x	16x
3 Kb Paired End	1.6x	1.8x	8x
20 Kb Paired End	1.6x	2x	1.5x

▲ **Table 1: Assemble Complex Genomes Using a Combination of GS FLX Titanium Series Multiple Span Paired End and Shotgun Reads.** The large scaffold size allows a detailed view of genome organization and enables straightforward finishing if desired.

## Go further, faster than ever before in complex genome assembly

- **One system. One solution:** The Genome Sequencer FLX System delivers all the tools necessary to sequence and assemble higher complexity genomes — long reads, long spans and sophisticated analysis software.
- **More complete genome coverage:** Long shotgun and 20 Kb span paired end reads, the longest of any next generation-sequencing system, ensure contiguous sequence information to cover the highly repetitive regions of complex genomes.
- **Comprehensive data:** Explore the full range of genetic variation including SNPs, indels, inversions, and rearrangements of any size.
- **Reduce project costs and time:** Replace the use of old, expensive and time consuming technologies such as fosmids and Sanger sequencing.
- **Minimize bioinformatics burden:** Compact data files eliminate the need for expensive computing and storage infrastructure, while the included analysis software significantly simplifies the assembly of even the most complex genomes.

# Highest Coverage. Longest Contigs. Fastest Assembly.

## GS FLX Titanium Series

### Multi Span Paired End Library Adaptors

Paired end reads (also known as “mate pair” reads) supplement shotgun reads to significantly improve *de novo* assemblies by completely or partially spanning highly repetitive genomic regions.

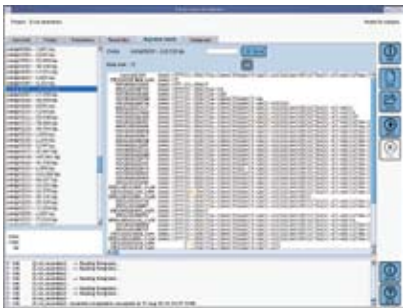
Paired end reads aid in:

- Ordering contigs into large scaffolds
- Determining the sequence within larger repeat regions
- Identifying structural variations
- Providing the complete genomic landscape, including gene order and operons



## Beyond Data to Insight

Easy-to-use software for quick and affordable assemblies



- The GS *De Novo* Assembler is a straightforward tool for automated *de novo* assembly of small to mid-size genomes (<1 Gb).
- Biologist or Bioinformatician - Benefit from an easy-to-use graphical user interface, as well as a powerful, scriptable command line interface.
- Modest system and expertise requirements eliminate the need for a high performance computing infrastructure and a dedicated bioinformatics team to produce results.
- Generate assemblies from a variety of shotgun and paired end reads - GS FLX Titanium series, GS FLX Standard series, GS 20 System, and Sanger reads.

Published by:

Roche Diagnostics GmbH  
Roche Applied Science  
68298 Mannheim  
Germany

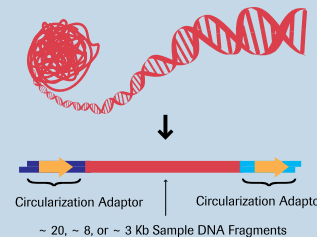
© 2009 Roche Diagnostics GmbH. All rights reserved.  
05588766001 • 4/09

For life science research only. 454, 454 SEQUENCING, GS 20 SYSTEM and GS FLX TITANIUM are trademarks of Roche.  
Other brands or product names are trademarks of their respective holders.

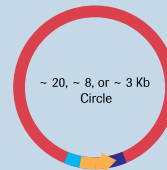
## The system's unique features include:

- **Long reads:** Extra long 400-500 base pair shotgun reads can sequence through many genomic repeat features directly without the need for paired ends.
- **Long tags:** Paired end reads with long tags, averaging 150+ bases, can be aligned uniquely with higher confidence.
- **Long spans:** Long 20 Kb paired end reads can span most repeat regions in nearly any size genome.
- **Span variety:** The broad selection of insert lengths enables optimization of project design according to the unique characteristics of any genome for the best possible results.

### The GS FLX Titanium Series Paired End Protocol



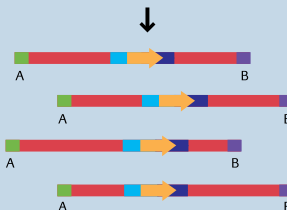
1. **Circularization-Ready Fragments.** Genomic DNA is sheared into 20 Kb, 8 Kb or 3 Kb fragments and adaptors are added to the end of each fragment.



2. **Circularized DNA.** DNA is circularized.



3. **Paired End Library Construct.** The circularized DNA is fragmented and fragments containing the added adaptors are isolated and amplified for sequencing.



4. **Paired End Library.** Resulting library consists of true paired end reads with two end tags averaging over 150 bp and separated by 20 Kb, 8 Kb or 3 Kb.

### Selected References:

**Rounsley S et al. (2009)** De novo next generation sequencing of plant genomes. *Rice*. 2(1): 35-43.

**Nagendran S et al. (2009)** Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus *Amanita bisporigera*, based on the secretome of *Trichoderma reesei*. *Fungal Genet. Biol.* 46(5):427-35.

Review the complete list of publications at [www.454.com](http://www.454.com).