



NGS Process Analysis – Market Trends, Unmet Needs & Challenges, and Ecosystem Analysis

Table of Contents – June 2016

Contents	2
Foreword	3
Learnings	4
The NGS Process	9
The published NGS literature	9
The Lab Components including Reagents & Sequencing Instruments	11
Top life science product suppliers in the NGS literature	11
Top products by reagents and sequencing instrument vendors	12
Sequencing instruments in the literature	14
Top 20 countries, institutions, and researchers across the NGS space	15
User type	16
The non-computational researcher	16
The computation researcher	16
End-user sample and library preparation reagents and kits preferences	18
End-user sequencing instrument preferences	20
The NGS Analysis and Interpretation Ecosystem	23
NGS utilization and analysis challenges	23
Secondary analysis companies addressing NGS utilization and analysis challenges	23
Tertiary interpretation companies addressing NGS utilization and analysis challenges	23
End-user feedback on secondary analysis	27
End-user feedback on tertiary analysis	39
Deep dive into the secondary analysis and tertiary interpretation tool companies	31
Secondary analysis tool companies	32
Data analysis speed, scalability, and flexibility	32
Sequencing machine integration	32
Tertiary interpretation tool companies	33
Tertiary interpretation content	33
A comparative functional assessment	36
Infrastructure requirements	36
Data security and compliance features	37
Partnerships	40
Overview of partnerships in the analysis and interpretation space	40
Company momentum	50
Concluding notes	52
Company/Product Profiles	57
Appendix	90
Companies in the data analysis and interpretation space	90
Top countries, institutions, and researcher in the NGS literature – data tables	93
About enlightenbio LLC	111

Sequencing reagents & instruments
(NGS Literature & end-user interviews)

NGS analysis & interpretation ecosystem analysis

Concluding notes with learnings

Company profiles - funding, # of employees, product details, company vision, and more

NGS Process Report – June 2016

The NGS process space is extremely busy with many commercial players that provide sequencing reagents, instruments, and analysis solutions. An extensive analysis of this space highlights that as the sequencing technology has matured, the challenges have shifted toward data explosion and computing, data interpretation to gather and report actionable information, and information management.

This report is unique, in that it is not a predictive market research report, but rather builds on data gathered from the many end-user interviews and the data available in the scientific literature.

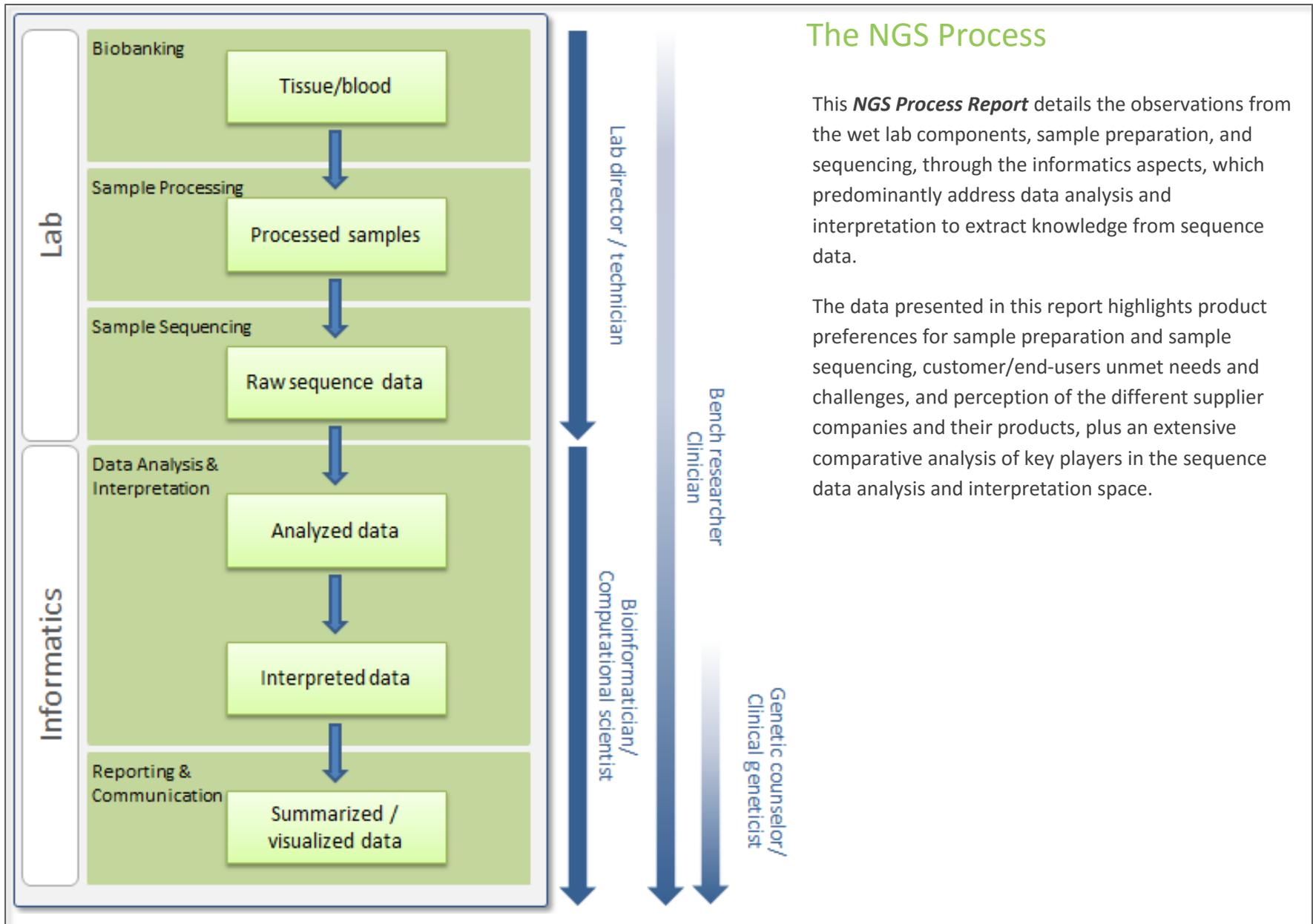
Our specific analysis identified the obvious top players which includes Illumina on the instruments side (and to some extent on the reagents side), Thermo Fisher Scientific, and Pacific Biosciences; Illumina also offers a data analysis and interpretation solution via BaseSpace. Agilent technologies, Qiagen, and Thermo Fisher Scientific are the main actors on the reagents side, while on the analysis side we observe several providers that are strong in different areas of data analysis and interpretation. DNAnexus and Seven Bridges Genomics have emerged as open and flexible platforms for data analysis, Bluebee, Genalice, and Edico Genome as alternatives enabling fast DNA processing. On the data interpretation side we have several noteworthy companies such as Congenica, Ingenuity (Qiagen), Omicia, Station X, or WuXi NextCODE offering elegant interpretation solutions. Alongside these established companies, we should not overlook the significant impact of open source tools that command a sizable share of the data analysis volume (~68% of all users still use in-house or open source tools for their analyses – some as standalone solutions and some in combination with commercially available offerings), such as the Broad Institute Scientific Community Software and Sanger Institute Tools and Software made available to the entire community.

Our data is substantiated with an **analysis of the published NGS/sequencing literature, 28 interviews (19 end-users and nine industry representative interviews)** highlighting product preferences for sample preparation and sample sequencing, customer/end-users unmet needs and challenges, and the brand perception users have for the different supplier companies and their products, plus **an extensive comparative analysis of key players in the sequence data analysis and interpretation space**. Furthermore, **this report provides lists of top global organizations and researchers** for the top sequencing applications (bioinformatics, RNA-seq, Sanger sequencing, exome, *de novo* assembly, variant calling, and CHIP-seq), as identified across the NGS/sequencing literature.

The approach taken to decipher the NGS process and its components, as summarized within this report, was rooted in the following questions: who are the key commercial companies in this process, who are the end-users and what are their unmet needs and challenges, and when researching the analysis and interpretation side of the equation, what challenges need the commercial life science product providers address, and who addresses them how?

This 111 page NGS Process Report consists of 23 Figures, 39 Tables, and 17 comprehensive company profiles (Agilent Technologies, Bina Technologies (Roche), Bluebee, CLC Bio (Qiagen), Congenica, DNAnexus, Edico Genome, Genalice, BaseSpace (Illumina), Ingenuity Systems (Qiagen), New England Biolabs, Omicia, Qiagen, Seven Bridges Genomics, Station X, Thermo Fisher Scientific, and WuXi NextCODE) which includes company metrics (funding, number of employees, etc.), product details, founder/executive and board information, additional notes, and the respective company vision.

For more information contact info@enlightenbio.com or visit our website, enlightenbio.com.



Foreword

This NGS Process Report details the observations from the wet lab components, sample preparation, and sequencing, through the informatics aspects, which predominantly address data analysis and interpretation to extract knowledge from sequence data. The data presented in this report highlights product preferences for sample preparation and sample sequencing, customer/end-users unmet needs and challenges, and the brand perception they have for the different supplier companies and their products, plus an extensive comparative analysis of key players in the sequence data analysis and interpretation space.

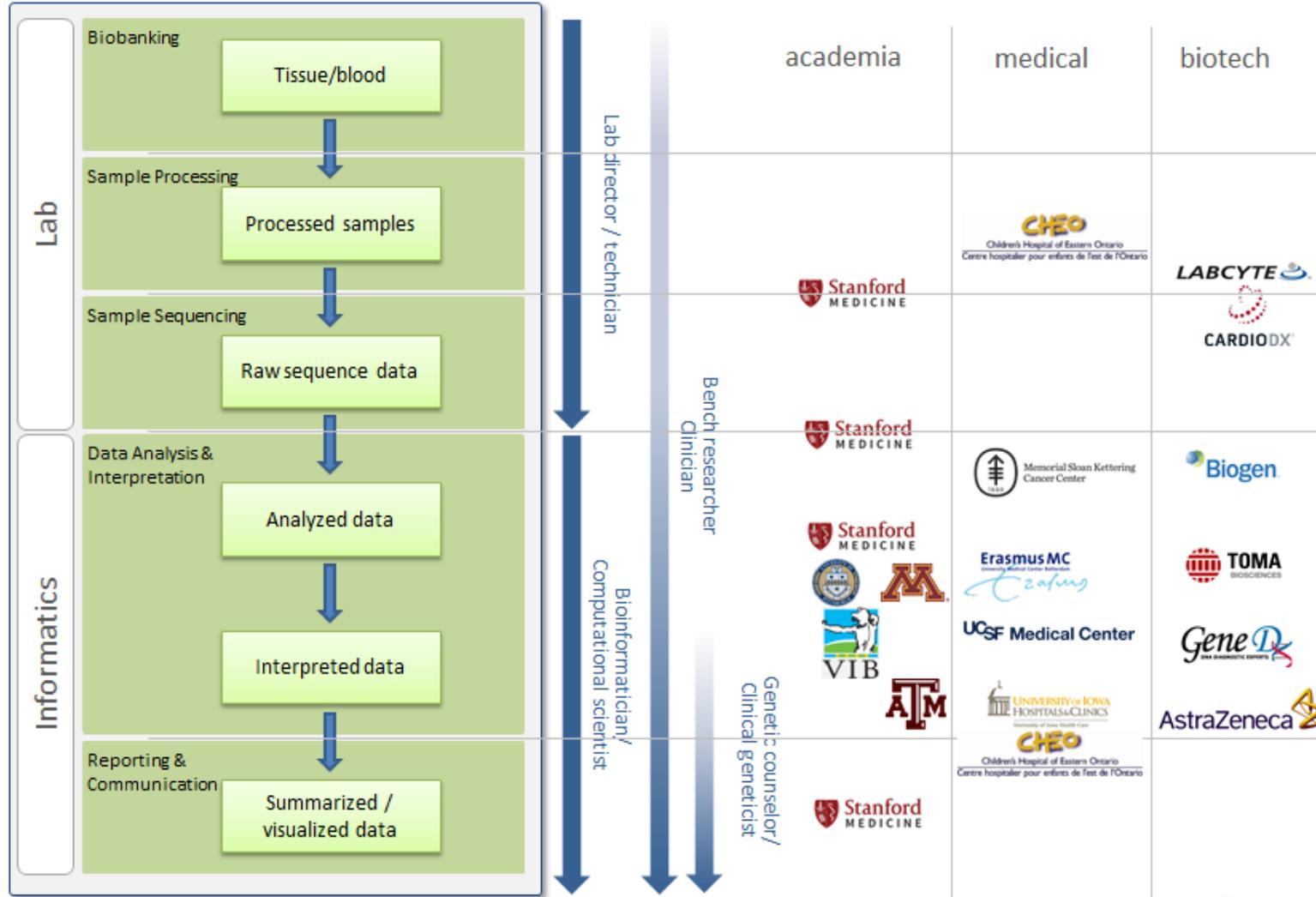
The approach taken to decipher the NGS process and its components was rooted in the following questions: who are the key commercial companies in this process; who are the end-users and what are their needs, pains, and challenges; and when researching the analysis and interpretation side of the equation, what challenges do these companies need to address, and are those challenges being properly addressed?

To create a robust report that answers these questions, we conducted our research following this investigative path:

- 1) Analysis of published literature data - contributed by our partner company [SeekQuence](#) - to understand NGS trends, commercial product supplier market share trends, and top performing NGS products and applications, among different countries, institutions, and researchers publishing in the space.
- 2) End-user interviews seek to understand user-processing needs, research product preferences, and challenges encountered throughout the NGS process.
- 3) Analysis of meta-data available via the World Wide Web to perform a deep dive interrogation of individual process components.
- 4) A deep ecosystem level Informatics analysis with a subset of key players in the secondary and tertiary analysis space, to understand what their focus is, how end-user needs are addressed, or what their company momentum is considering their funding situation, number of employees, and more.
- 5) Company profiles for the key players across the entire NGS process.
- 6) Input from key representatives for established commercial research product suppliers to learn about their vision for product solutions in the current - and future - space.

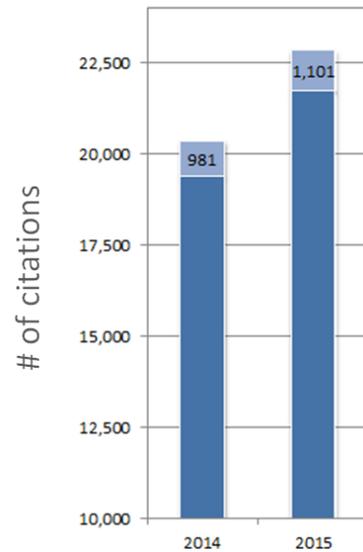


End-User Interviews



19 end-user interviews conducted to elucidate processing needs and challenges, and research product preferences.

Published NGS Literature Analysis



An *analysis of published literature data*

Contributed by our partner company [SeekSequence](#) this section aims at understanding NGS trends, commercial product supplier market share trends, and top performing NGS products and applications, among different countries, institutions, and researchers.

- Term "Sanger sequencing"
- Term "sequencing"

Top 20 countries, institutions, and researchers across the NGS space

A series of citation analyses were performed to identify the leading countries, institutions, and dominant researchers for the top applications as summarized in Figure 3. In all cases, the identified 2015 NGS publications underwent a full text search for the individual terms that define specific NGS applications. Each table contains: A) the top countries, B) the top institutions, C) the top institutions based on published data in journals with an impact factor of 5 or greater (≥ 5), and D) the top researchers. The individual data tables can be found in the appendix.

Application researched	Table number	Page
Bioinformatics	Table 34	96
RNA-seq	Table 35	98
Sanger sequencing	Table 36	100
Exome	Table 37	102
<i>de novo</i> assembly	Table 38	104
Variant calling	Table 39	106
ChIP-seq	Table 40	108

Table 1: Overview table NGS application research

The NGS Analysis and Interpretation Ecosystem

Deep dive into secondary analysis and tertiary interpretational tool companies

Numerous companies exist in the secondary and tertiary analysis space – a non-exhaustive list can be found in the Appendix in Table 31. Many of these companies are either niche companies, only tangential to the secondary or tertiary analysis space, or do not have the same company momentum as the companies included in this analysis. This analysis focuses on 13 interesting current players (see the table to the right) with high company momentum. The companies selected have high funding power, a well-received customer perception, are strategically well connected with relevant partner companies, or have in general high visibility in the sector. Not included are companies that focus on information management or clinical reporting.

2°	Bluebee
	CLC bio/Qiagen
	DNAexus
	Edico Genome
	Genallice
2°/3°	Seven Bridges
	BaseSpace/Illumina
	Bina Technologies/Roche
3°	Congenica
	Ingenuity/Qiagen
	Omicia
	Station X
	WuXi NextCODE



17 Company/Product Profiles

- Agilent Technologies
- Bina Technologies (Roche)
- Bluebee
- CLC Bio (Qiagen)
- Congenica
- DNAnexus
- Edico Genome
- Genalice
- BaseSpace (Illumina)
- Ingenuity Systems - IPA & IVA (Qiagen)
- New England Biolabs
- Omicia
- Qiagen
- Seven Bridges Genomics
- Station X
- Thermo Fisher Scientific
- WuXi NextCODE

Company/product profiles of life science product and solution providers that supply relevant reagents or instruments in the sequencing sector, or provide data analysis and interpretation software (and hardware) included in the deep dive ecosystem analysis.

Company profiles highlight company metrics (funding, number of employees, etc.), product details, founder/executive and board information, additional notes, and the respective company vision.

Companies mentioned in report

10xGenomics, Affymetrix, Agilent Technologies, Amazon, Appistry, Beckman Coulter, Bina Technologies, Biomatters, Bluebee, Cartagenia, CLC Bio, Complete Genomics, Clontech, Congenica, Core Informatics, DNAnexus, DNASTAR, Enlis Genomics, Edico Genome, Epicenter, Fluidigm, Genia (Roche Sequencing System), Genalice, GeneData, GenomOncology, GenoLogic (Illumina), GenoSpace, Golden Helix, Google, Illumina, IBM, IMS Health, Ingenuity Systems, Lab7 Systems, Kapa Biosystems, Microsoft, Maverix Biomics, Nanostring, New England Biolabs, N-of-One, Nugen, Oxford Nanopore, Pacific Biosciences, Partek, Perkin Elmer, Qiagen, Raindance, Roche, Rubicon Genomics, Sapio Sciences, Seven Bridges Genomics, SolveBio, SoftGenetics, Spiral Genetics, Station X, Swift Biosciences, Thermo Fisher Scientific, Tute Genomics, WuXi NextCODE Genomics.

About enlightenbio LLC

Enlightenbio was founded in 2013 in the San Francisco Bay Area to provide a conduit between research and related technical and analytical resources. Our company consists of PhD level research scientists who bring decades of industry experience and expertise in the biotechnology, molecular diagnostics, pharma, and life science research markets to the table. We are dedicated to communicating in the researcher's language, identifying unmet needs, and understanding product development. Our goals are aligned with researcher's needs to increase experiment productivity and to make sense of the resulting biological data.

In addition to our varied industry experiences - Applied Biosystems (now Thermo Fisher Scientific), Iconix Biosciences, Ingenuity Systems (now a Qiagen company), and DNAnexus – we have built and maintained content curation services, defined product strategy, managed tactical product projects, performed extensive ecosystem analyses, and defined go-to-market plans.

Building on our initial success and previous experiences - microarray and next-generation sequence data analysis, toxicogenomics, solutions for sequence data management, analysis, and interpretation, drug discovery, and biochemistry - we continuously monitor worldwide market trends in healthcare information technology, life sciences, genomics, clinical diagnostics, and in the medical devices space to expand our critical service offerings. Combined with our extensive global network, we can identify target market pain points and unmet needs, perform detailed market and product research, perform horizontal and vertical ecosystem or competitive analyses, and more.

This background with our future-focused outlook makes us a resourceful and exciting alternative to traditional market research companies. Our comprehensive knowledge of the market we live and breathe is invaluable to our partnerships and to the potential for our customers.

enlightenbio and associated market research reports are managed by Brigitte Ganter, PhD, Founder & Managing Director of enlightenbio LLC.