

# Next-Gen Sequencing Trends 2009



August 2009

[www.htstec.com](http://www.htstec.com)

## Conditions Under Which This Market Report Is Sold

**This REPORT is Copyright protected by HTStec Limited. All rights reserved. Purchase of an electronic license to this REPORT entitles you to use it solely and exclusively within the purchasing Company. Neither this REPORT nor any of its contents may be disclosed or transferred by any means (electronic or otherwise) to ANY third party (i.e. beyond the purchasing Company) without the prior written approval of HTStec Limited.**

HTStec Limited has exercised due care in compiling and preparing this REPORT, which is based on information submitted by individuals in respondent companies. HTStec Limited has NOT verified the accuracy of this information, nor has it established respondent's authority to disclose information to HTStec Limited. HTSTEC LIMITED EXPRESSLY DISCLAIMS ANY AND ALL WARRANTIES CONCERNING THIS REPORT, INCLUDING ANY WARRANTIES OF MERCHANTABILITY AND/OR FITNESS FOR ANY PARTICULAR PURPOSE, AND WARRANTIES OF PERFORMANCE, AND ANY WARRANTY THAT MIGHT OTHERWISE ARISE FROM COURSE OF DEALING OR USAGE OF TRADE. NO WARRANTY IS EITHER EXPRESSED OR IMPLIED WITH RESPECT TO THE USE OF THE REPORT. Under no circumstances shall HTStec Limited be liable for incidental, special, indirect, direct or consequential damages or loss of profits, interruption of business, or related expenses that may arise from use of this REPORT, including but not limited to those resulting from inaccuracy of the data therein.

## Executive Summary

- This market report summarizes the results of HTStec's industry-wide global web-based benchmarking survey on next-generation (next-gen) sequencing carried out in July 2009.
- The study was initiated by HTStec as part of its ongoing tracking of emerging life science technologies and marketplaces. The main objectives were to document current user experiences, practices, assay metrics and preferences obtained using next-gen sequencing instruments and to understand future requirements. Equal emphasis was given to soliciting opinion from all organizations where next-gen sequencing is being undertaken, with no geographic bias in the distribution of persons contacted.
- The survey looked at the following aspects of next-gen sequencing as practiced today (2009) and in some cases as predicted for the future (2011): sequencing platforms respondents use or have access to; applications run on sequencing platforms; current bottlenecks; main bottleneck in sample prep; sample prep methods used to shear DNA; aspect of sample prep most in need of automation; typical cost per mappable base sequenced; mappable bases typically read per run; sequencing throughput achieved; biggest restriction to improving sequencing throughput; how frequently systems have broken down; experience of vendor service support; whether any publishable data has been generated; platforms most associated with desirable characteristics/attributes; what is used for data analysis and storage; biggest bottleneck in data analysis and storage; where respondents plan future growth in data analysis and storage; how long raw data is retained; future expectations for cost per genome; improvements most wanted in next-gen sequencers; whether a 'gold standard method' will emerge; thoughts on achieving a \$1,000 genome; what has impacted respondent's ability to purchase/implement a next-gen sequencing capability; plans to purchase new next-gen sequencing systems; thoughts on 3<sup>rd</sup> generation sequencing solutions; interest in outsourcing next-gen sequencing, providers used/considered; budget allocated and its breakdown; and unmet needs in the next-gen sequencing process.
- The main questionnaire consisted of 29 multi-choice questions and 2 open-ended questions. In addition, there were 5 questions related solely to the administration and demographics of the survey.
- The survey collected 172 responses. Approximately 40% (68) of these responses were subsequently excluded as they did not meet the initial respondent selection criteria, or provided minimal input, or came from vendor companies active in the area. This left 106 survey respondents that met our selection criteria (i.e. were currently undertaking or involved in next-gen sequencing), provided verifiable details of their lab origin and gave input in a significant part of the questionnaire.
- Survey responses were geographically split: 53% Europe; 27% North America; 11% Asia (Excluding Japan); 5% Japan; and 4% Rest of World.
- Respondents came from 72 University, Research Institute or Not-for-Profit Facilities; 10 Biotechs; 7 Other; 5 Not-For-Profit Sequencing Centers; 5 Government/Military/Defense; 4 Commercial Fee-For-Service Providers; 1 Pharma; 1 Diagnostics Company; and 1 Hospital/Clinic.
- Most survey respondents had a senior job role or position which was in descending order: 15 Lab/Research/Facilities Manager; 15 Professor/Assistant Professor; 13 Research Scientist; 11 Department Head; 8 Principal Investigator; 7 Director; 6 Section/Group Leader; 5 Senior Scientist/Research Associate; 4 Post-doc; 3 Vice President; and 19 Other Roles.
- Survey results were expressed as an average of all survey respondents. In addition, where appropriate the data was reanalyzed after sub-division into the following 4 survey groups: 1) University or Research Institute; 2) Other Organizations; 3) Europe; and 4) North America.
- Survey respondents use in their labs or have access to (e.g. via collaborating centers) the following next-gen sequencing platforms: 82% to 454 FLX/Roche, 79% to Solexa/Illumina, and 56% to SOLiD/ABI.
- The application most run on a next-gen sequencer was mutation/SNP discovery.
- Data processing was the most limiting bottleneck in next-gen sequencing today.
- The main bottleneck in sample preparation for next-gen sequencing reported was library construction.
- The most used method for DNA shear in next-gen sequencing sample prep was nebuliser for both short and medium/long fragments, and sonication for fragmentation for ChIP.
- The aspect of sample prep for next-gen sequencing most in need of automation was DNA fragmentation.
- The median typical cost per billion mappable bases sequenced was <\$25K for 454 FLX/Roche; <\$2.5K for SOLiD/ABI; and <\$2.5K for Solexa/Illumina.

- The median typical number of mappable bases read per run today was <0.5 billion for 454 FLX/Roche; <5 billion for SOLiD/ABI; and <5 billion for Solexa/Illumina.
- The median throughput reported for next-gen sequencing operations today was <0.5Gb/day for 454 FLX/Roche; <1Gb/day for SOLiD/ABI; and <1Gb/day billion for Solexa/Illumina.
- Data analysis (quality of software tools) was rated the biggest restriction to improving throughput.
- The highest incidence of frequent breakdowns was reported for Solexa/Illumina.
- The median time interval between system failures on all next-gen sequencing platforms was 1–3months.
- The service support experience related to 59 next-gen sequencing platforms used by survey respondents are reported with respect to overall satisfaction, faults rectified on first visit and geographic location.
- The highest incidence of very happy with overall service support experience was reported for SOLiD/ABI.
- The highest incidence of >99% of faults rectified on first visit was reported for 454 FLX/Roche.
- The highest incidence of generating publishable data was reported for 454 FLX/Roche.
- The next-gen sequencing platform most associated with the following desirable characteristics/attributes was: 454 FLX/Roche for system reliability, ease of integration with LIMS, and understanding sequencing/user requirements; SOLiD/ABI for good after-sales support & service, and innovative software; and Solexa/Illumina for highest accuracy, cutting-edge technology, and value for money.
- The majority of respondents used in house bioinformatics for data analysis and storage.
- The biggest bottleneck for next-gen sequencing data analysis and storage was software.
- The median time respondents plan to retain raw data (unprocessed image files) was 3–6 months.
- The future requirements (i.e. where they expect be by the end of 2011) of 52 respondents for the cost per genome sequenced are reported with respect to base pair read length (fragment size) processed.
- The median cost per genome most expected was <\$10K for short fragment applications of >100kb and long fragment applications of 1.5–3kb.
- Read length accuracy (irrespective of the definition) was rated what respondents most wanted to see improved in the next-gen sequencing offerings.
- Most respondents do not believe that any of the 3 established next-gen sequencing platforms will become *'the gold-standard method'* in the next few years.
- The majority of respondents believe a \$1,000 genome is achievable, within a median time frame of <3 years, and will most likely to be achieved using a PacBio platform.
- Infrastructure requirements were ranked what had most impacted respondent's ability to purchase and/or implement a next-gen sequencing capability in their organization.
- A total of 62 purchasing plans for next-gen sequencers spread over the next 3 years were identified from 32 different labs. Greatest interest was for purchasing Solexa/Illumina. The majority of labs planned to purchase more than 1 unit, mainly from a single instrument manufacturer.
- Most respondents are not delaying initiating sequencing and putting off the purchase of a sequencing platform until 3<sup>rd</sup> generation sequencers come to the market.
- Opinion on 3<sup>rd</sup> generation sequencers was that the need for instrument upgradeability is absolutely essential and the importance of a common infrastructure is desirable.
- Most respondents are currently outsourcing at a not-for-profit collaborating lab, only a small minority use commercial fee-for-service providers, and a list of the providers used to date was collated.
- The median outsourcing budget for next-gen sequencing at a fee-for-service provider was \$10K–\$25K.
- The breakdown of the budget for outsourced next-gen sequencing at a fee-for-service provider was projected with the biggest proportion spent on transcriptome analysis, small RNA analysis and mutation/SNP discovery.
- A simple bottom-up model was made to estimate the market for outsourced next-gen sequencing at a fee-for-service provider. This model used data derived from this survey on respondent's involvement in outsourcing and their mean budgets allocated or estimated per lab for this activity.
- The global 2009 market estimate for outsourced next-gen sequencing at a fee-for-service provider was around \$30million. The full report details the 2011 estimate, % CAGR, and its breakdown into services.
- Respondent's feedback on the biggest unmet needs in next-gen sequencing was documented.
- The full report provides the data, details of the breakdown of the responses for each question, mean and median values for the assay metrics, and the estimates for the future (2011). It also highlights a few interesting differences between the survey groups.

## Table of Contents

Executive Summary .....	2
Table of Contents .....	4
Survey Methodology .....	5
Respondent's Company or Organisational Origin .....	7
Respondent's Geographic Origin.....	8
Respondent's Job Role .....	9
Next-Gen Sequencers Used or Accessed.....	10
Application Run on a Next-Gen Sequencer .....	11
Current Bottleneck in Next-Gen Sequencing .....	12
Main Bottlenecks in Sample Prep for Next-Gen Sequencing.....	13
Sample Prep Methods Used for DNA Shear.....	14
Aspect of Sample Prep in Need of Automation .....	15
Cost Per Mappable Bases .....	16
Mappable Bases Read Per Run (1).....	17
Mappable Bases Read Per Run (2).....	18
Throughput – Mappable Bases Read Per 24h day (1) .....	19
Throughput – Mappable Bases Read Per 24h day (2) .....	20
Biggest Restriction to Improving Sequencing Throughput .....	21
Summary of Survey Findings (1).....	22
Frequency of Breakdown of Next-Gen Sequencing Systems.....	23
Time Between Failures in Next-Gen Sequencers .....	24
Vendor Service Support – Overall Experience (1) .....	25
Vendor Service Support – Overall Experience (2).....	26
Vendor Service Support – Fault Rectification .....	27
Vendor Service Support – Geographic Location of Systems Surveyed .....	27
Generated Publishable Data .....	28
Next-Gen Sequencing Solution Most Associated With Desirable Attributes .....	29
What is Mainly Used for Next-Gen Sequencing Data Analysis & Storage .....	30
Biggest Bottleneck in Next-Gen Sequencing Data Analysis & Storage.....	31
Where Respondents Plan Future Growth in Next-Gen Sequencing Data Analysis & Storage.....	32
How Long Respondents Plan to Retain Raw Data .....	33
Cost Per Genome – Future Requirements for Next-Gen Sequencing (1).....	34
Cost Per Genome – Future Requirements for Next-Gen Sequencing (2).....	35
Cost Per Genome – Future Requirements For Next-Gen Sequencing (3) .....	36
Improvements Wanted in Next-Gen Sequencing Offerings .....	37
Thoughts on a Gold Standard Method for Next-Gen Sequencing.....	38
Time Frame to Achieve a £1,000 Genome.....	39
Likely Platform to Achieve a \$1,000 Genome .....	40
Summary of Survey Findings (2).....	41
What Has Impacted Ability to Purchase a Next-Gen Sequencing Capability .....	42
Plans to Purchase Next-Gen Sequencers (1).....	43
Plans to Purchase Next-Gen Sequencers (2).....	44
Respondents Waiting for 3 <sup>rd</sup> Generation Sequencers.....	44
Opinion on 3 <sup>rd</sup> Generation Sequencers.....	45
Current Level of Interest in Outsourcing .....	46
Fee-For-Service Providers Used .....	46
Annual Budget for Outsourced Next-Gen Sequencing .....	47
Breakdown of Outsourcing Budget for Next-Gen Sequencing .....	48
Samples Analysed per Year Using Outsourced Genotyping .....	49
Biggest Unmet Needs in Next-Gen Sequencing.....	50
Summary of Survey Findings (3) .....	51

## General Information on HTStec and HTStec's Trends Market Reports

- HTStec Limited an independent market research consultancy founded in September 2003 whose focus is on assisting clients delivering novel enabling platform technologies (liquid handling, laboratory automation, detection instrumentation and assay reagent technologies) to drug discovery and the life sciences. Over the past 6 years HTStec has published 48 market reports mainly on drug discovery technologies and authored 29 technology review articles in Drug Discovery World.
- HTStec's Trends reports owe their origins to the need by developers and vendors of new enabling technologies in drug discovery to get up-to-date relevant market metrics on which to base informed business decisions.
- Typically focused on a specific market niche or segment, in many cases overlooked or frequently misunderstood by broader market studies.
- Investigations are mainly initiated in response to a sponsor's specific requests.
- HTStec's extensive experience of the market, both as a Pharma End-User and working for a major Life Science Tool Provider ensures the industry relevance of the market research collected.
- Based entirely on web-based feedback from potential customers typically drawn mainly from Pharma and Biotechs, although increasingly University and Research Institute labs are also being researched.
- Produced extremely rapidly and typically published within 3 weeks of starting the collection phase.
- Reports are short, concise and focused on giving readers the basic data, analyzed in several different ways.
- Limited to reporting the main findings alone, without exhaustive discussion on the relevance of the results.
- Market estimates are mainly based on bottom-up calculations and usually avoid attempts to forecast widely beyond the next 2-3 years. Full details on the derivation of market estimates are given so readers can apply their own factors and easily make alternative estimates if they prefer or know better.
- Owing to the sensitivity of some of the data collected, all reference to the origin of participating companies is removed, data is pooled to get an industry average and the anonymity of all respondents fully preserved and guaranteed.
- Unlike alternatives HTStec's Market Surveys and Report are aimed at giving readers, information they want and can rely on, not information they don't need, cannot easily discern or is of dubious authenticity.
- HTStec aims to be the premier global provider of highly focused market research on enabling technologies in drug discovery.
- To get information or to request free executive summaries of published reports please contact [john.comley@htstec.com](mailto:john.comley@htstec.com).
- HTStec Limited is a privately owned UK Company, registered in England and Wales Number 4875933.