An expedited workflow for identification of bacteria by Next Generation DNA Sequencing: ultra-fast sample prep with streamlined data analysis

Dara S Missan [2]; Jeremy Ellis [2]; Paul Reeves [1]; Jan Rogers [1]; Pablo Castan [1]; Beatriz Lopez-Quintana [3]; Alicia de Pablo [3]

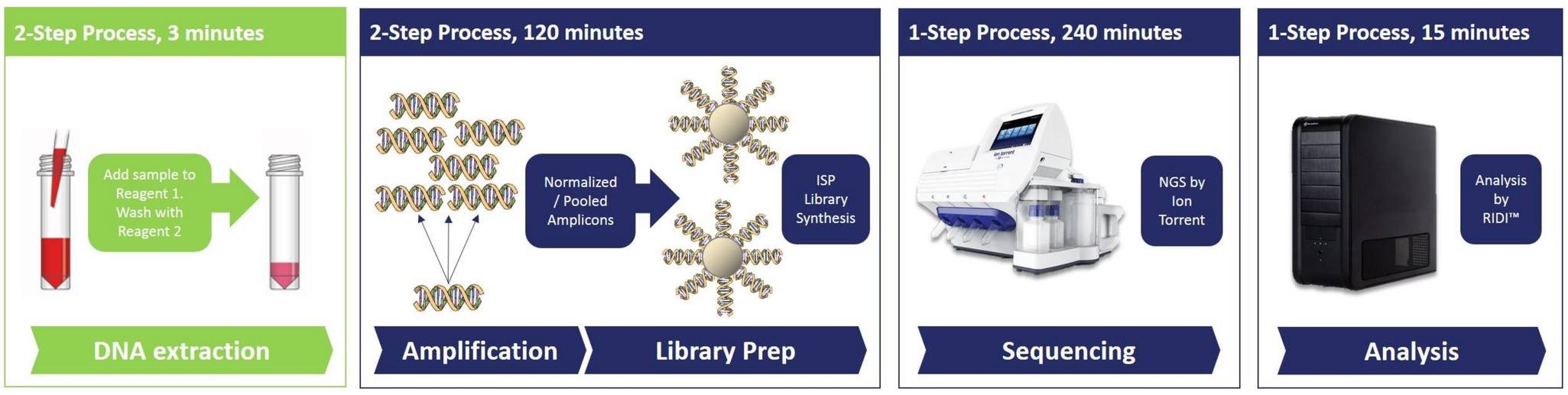
[1] Arcis Biotechnology (www.arcisbio.com), Innovation Centre, Sci-Tech Daresbury, Warrington, WA44FS, UK. [2] Fry Laboratories LLC (https://frylabs.com), 14807 N 73rd St Suite #103, Scottsdale, AZ 85260, USA. [3] Hospital Carlos III; Laboratorio de Micobacterias y Microbiologia Molecular, Madrid, Spain

INTRODUCTION

Clinical diagnostics based on Next Generation DNA Sequencing (NGS) technologies are gaining momentum as part of routine clinical practice. While offering considerable advantages for the simultaneous interrogation of multiple targets, the process of going from sample to sequence, including subsequent data analysis, is currently laborious with an unfavorable time to result. For clinical diagnosis of infectious disease there is a need to expedite the workflow in order to provide rapid identification of pathogens and inform treatment strategy. Indeed, in the case of sepsis, time to appropriate antimicrobial therapy is an independent predictor for death.

METHODS

In an effort to accelerate the clinical NGS workflow, two technologies were investigated for their co-compatibility and reduction of time-to-result; a 3 minute sample prep method (Arcis Biotechnology) was compared with a market leading column based DNA extraction kit. A rapid data analysis system was used for bacterial identification (RIDI, Fry Laboratories). Sequences were generated on the Ion Torrent PGM platform (Life Technologies). The entire workflow, from a single sample to pathogen identification was approximately 6.5 hours, for multiple samples the workflow was less than 8 hours.



RESULTS AND DISCUSSION

Escherichia Genus Amplicon Detection Rates

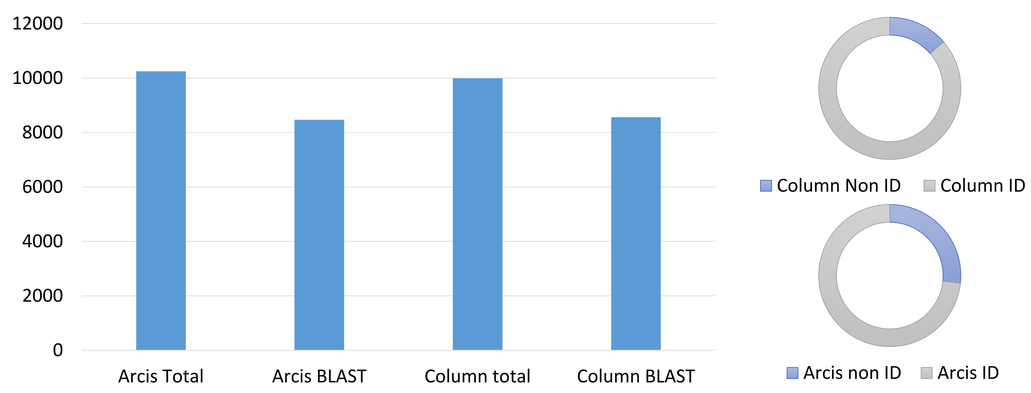


Figure 1. Similar performance observed between column extraction and Arcis extraction (E. coli spiked into blood). The graph shows the total number of sequences generated, and the number with valid BLAST returns. The pie charts indicate how many sequences match the target species within 95% or greater. Processing time for the column based kit was 60 minutes, processing time for the Arcis kit was 3 min.

Comparison of Extraction Methods on Proficiency Panel

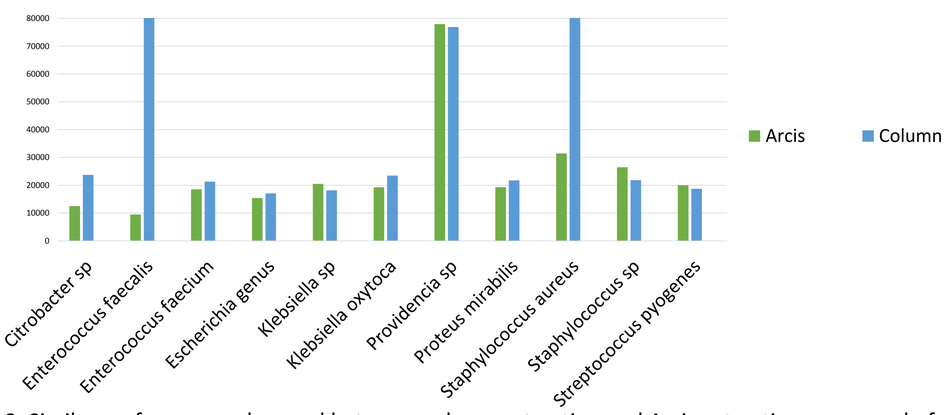


Figure 2. Similar performance observed between column extraction and Arcis extraction on a panel of proficiency testing samples, representing a range of Gram positive and Gram negative bacteria. Two outliers to the data were Enterococcus faecalis and Staphylococcus aureus, which gave considerably more sequences on the column-based extraction system. The RIDI data analysis software rapidly performed all data processing steps including sequencing filtering, trimming and editing.

Proficiency Samples and Detection Robustness

Sample ID	Template Vol (µl)	Rxn Vol (µl)	Total Seqs	BLAST Seqs	%ID
Proteus sp	1	20	24096	15134	56.9
	5	50	34764	23527	54.8
Citrobacter sp	1	20	26194	21655	95.2
	5	50	28286	21491	93.3
Klebsiella sp	1	20	33123	26586	89.5
	5	50	43630	35678	93.0
Staphylococcus sp	1	20	33081	26219	89.6
	5	50	37432	30057	91.8
Enterococcus faecium	1	20	20795	10817	95.3
	5	50	21508	12777	94.4

Table 1. PCR Protocol Optimisation.

Template vs Reaction Volume Ratios For Escherichia coli

Template Vol (µl)	Rxn Vol (µl)	Template:Rxn	Total Seqs	BLAST Seqs	% ID
50	150	0.33	28708	25362	98.9
25	75	0.33	30841	15695	96.9
10	30	0.33	4714	4313	96.8
1	20	0.05	4354	4002	99.1

Table 2. PCR Protocol Optimisation. An assessment of input template volume and total reaction volume suggests that the protocol is robust with good performance across a range of input values. Once a set of trimmed reads is obtained from the RIDI software, the sequences are identified using a unique reductive BLAST strategy. The % ID results, generated on E.coli spiked blood, were greater than 96% regardless of protocol used.

CONCLUSIONS

The results suggest good compatibility between the technologies, promising results in terms of sequences generated and identified. Sample prep time was reduced from 60 minutes to 3 minutes. Rapid sample prep combined with RIDI data analysis considerable allows time savings in a clinical NGS workflow.