NIIMBL Adventitious Agent Detection by NGS

AccuGenomics: Tom Morrison (PI), Bradley Austermiller, Nick Lazaridis, Chris Holshouser; *North Carolina State University*: Caroline Smith-Moore (CPI), Karen O'Connell; *Celgene*: Peter Bernhardt (CPI); *Merck*: Veronica Fowler (CPI); *NIIMBL*: Melissa Scott (PL), Peggy Thomas, Nicole Hoover

Background

This project consists of the development of an innovative testing platform, the AccuKit™ BioContaminants, which streamlines adventitious agent (AA) screening while directly ensuring every target is measured with the required sensitivity needed to pass Quality Control benchmarks. AccuKit BioContaminants incorporates PCR primers specific for 19 known adventitious viruses for use in high throughput next generation sequencing (NGS) screening assays. Our kit contains a mixture of competitive templates for every adventitious agent spiked into every sample as internal sensitivity controls; the expected test result for non-contamination is positive for every spike-in target and negative for the native targets. Importantly, the platform allows for additional primer sets to be added to the AccuKit BioContaminants to meet future adventitious agent detection needs. AccuKits increase the efficiency of adventitious agent detection by providing the missing quality controls to ensure a sensitive biosafety testing package for any biopharmaceutical product. This poster discusses the workflow and completion of the NIIMBL grant milestones.

Milestone 1: Primers and Controls for at least 15 agents

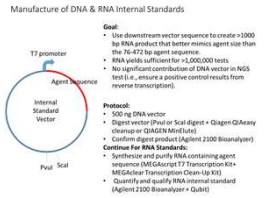
The goal of milestone 1 was to manufacture primers and controls for at least 15 adventitious agents (AA) with ten copy sensitivity. A list of 102 AA was whittled down to 24 (table below) due to budgetary concerns based on input from the Tier 1 partners Celgene, Merck, and NCSU. AccuGenomics objective was to create a highly sensitive and streamlined NGS testing process compatible with sample to **test results in less than a day**. The NGS library preparation was designed to detect as low as ten copies AA from a sample matrix containing 1ug DNA + 1ug RNA. 132 primers were selected from the literature and were screened for compatibility in NGS library preparation. Several primers were eliminated bioinformatically due to poor homology during blast searching. Other primers were ranked based on amplicon size and number of degenerate bases and were pooled and subjected to multiplex PCR followed by measuring yield by individual SYBR qPCR detection. Primers were then graded based on correct size, sensitivity, and no template control performance. Multiple rounds of pool combinations were required to narrow down the list to the final three primer pools capable of detecting 19 agents.

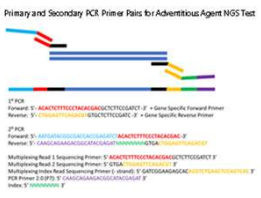
Initial Ranked Adventitious Agent List

Agent	Rank	Alias	Identified in Literature	Bioinformatically Eliminated	Screened	Backup	Included in AccuKit
Minute virus of Mice	1	MVM	7		6	1	1
Bovine Polyoma Virus	2	BPV	7	1	3	3	1
Cache Valley Virus	3	CVV	11	3	6	2	1
Reovirus 1/2/3	4	REO	14	1	4	9	0
Vesivirus 2117	5	VV	9	5	4	0	1
Epizootic Haemorrhagic Disease virus	6	EHDV	8		4	4	1
Rabies Virus	7	RAB	5		3	2	0
Bovine Circovirus	8		0				1*
Porcine Circovirus 1	9	PCV	3		3		1
Porcine Parvovirus	10	PPAV	4		4		0
Bovine Adenovirus	11	BADV	9	2	6	1	1
Bovine Parvovirus	12	BPAV	7		3	4	1
Bluetongue Virus	13	BTV	7	2	3	2	1
Bovine Coronavirus	14	BCOV	4		3	1	1
Bovine Enterovirus	15	BEV	4		4		1
Bovine Respiratory Syncytial Virus	16	BRSV	3		3		1
Porcine Adenovirus	17	PADV	4		4		0
Porcine Transmissible Gastroenteritis Virus	18	PTGV	5	1	4		1
Porcine Hemagglutinating Encephalomyelitis Virus	19	PHEV	4		4		1
Human Adenovirus 2	20	HADV	4	1	3		1
Bovine Parainfluenzavirus 3	21	BPIV	3		3		1
Porcine Circovirus 2	22	PCOV2	4		3	1	1
Bovine Viral Diarrhoea Virus	23	BVDV	3		3		1
Seneca Valley Virus (Senecavirus)	24	SVV	3		3		1

* Bovine Circovirus can be detected by the Porcine Circovirus 2 primer (PCOV2)

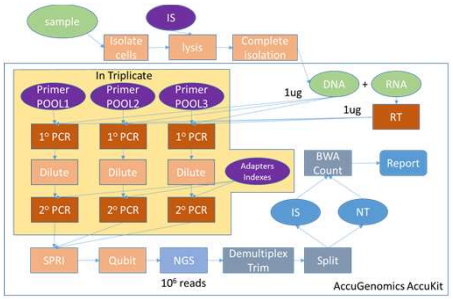
The next step in the project was to create an internal standard mixture at limit of detection levels (10 copies of each control into primary PCR). Internal standards were designed and ordered for the 19 AA as synthetic templates detectable by NGS (left figure below). To distinguished the controls from native AA, two adjacent bases in conserved AA regions were made. The two base changes were made in a region that would also support Taqman SNP based exonuclease hydrolysis probe-based detection as an alternative to NGS detection. For the RNA viral AA standards, in vitro transcription from an upstream t7 promotor would generate an RNA molecule which included an additional 1100 bases vector sequence. Like the DNA AA controls, the idea of including adjacent plasmid sequences was to generate a nucleic acid length more like the native viral sequence in an attempt to better mimic nucleic acid purification yields. As the RNA controls were derived from DNA templates, we tested each RNA control in absence of reverse transcription to ensure no significant DNA template remained (data not shown). An accelerated stability study for four of the RNA controls indicated that the control concentration would drop <25% over 3 years if stored at -80C (data not shown). Based on Qubit measurements, a master mixture of RNA and DNA standards were diluted to a mixture of 1000 copies each /ul. Two independent master stock mixtures were created from concentrated stocks (1010-1012 copies/ul) and measured by NGS as an estimate of manufacturing reproducibility. No obvious problems with reproducibly generating 1000 copies/ul master stocks from concentrated stocks were detected (data not shown).





Library Preparation and Sequencing Pipeline

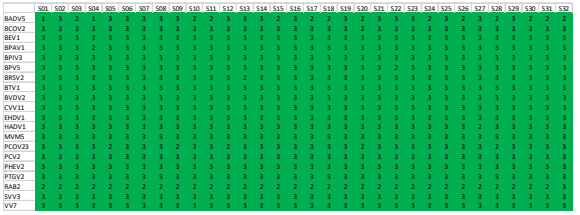
Cell lysates are spiked with DNA and RNA IS controls then co-purified. The isolated RNA is reverse-transcribed (RT) to cDNA then added to a primary PCR along with the isolated DNA. Primary PCR is divided into three separate reactions to ensure the resulting products are of similar yield as to not dominate the sequencing chemistry. Following the primary PCR, reactions are diluted and then added to a secondary PCR which adds the sequencing adapters and sample indices. Finally reactions are cleaned using SPRIselect beads to remove remaining primers, incomplete products and enzymes which may interfere with sequencing. The cleaned libraries are quantified, diluted and combined with PhiX, then sequenced on the Illumina MiSeq using single end sequencing with 75 cycles. The sequencing reads are de-multiplexed, trimmed, and split into native and internal standard bins. The data is imported into Excel where a summary table that contains the detection count for both NT and IS is generated for each sample. The entire process from sample to test results for ten samples takes less than a day.



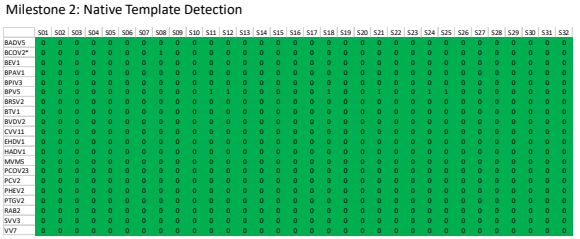
Milestone 2: Detection of IS in native template negative samples

Milestone 2 stated all of the agent negative specimen will not fail sensitivity Quality Control. To test this 32 negative samples were created so that 10 copies of each internal standards were present in each of the three primary PCR reactions. Due to the nature of sequencing, false positive reads can arise. To account for this, we have set an limit of blank (LOB) of 4 read counts, which is six standard deviations above background. Sensitivity was sufficient for each agent if at least 1 of 3 sample replicates have internal standard counts above LOB. An AA is positive in a sample when at least 2 of 3 sample replicates are above LOB. We tested the reagents created in Milestone 1 on 32 samples in triplicate and determined that 19/19 AA measurements for each sample met the requirement of IS detection and no detection of NT (figure below, rows are AA target ID, columns are sample ID, numbers indicate replicates above LOB).

Milestone 2: Internal Standard Detection



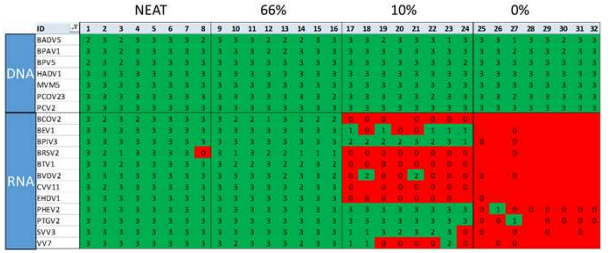
Milestone 2: Native Template Detection



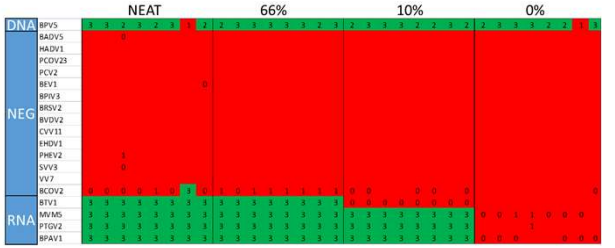
Milestone 3: Response of test to inhibition

Milestone 3 states that "When inhibiting reverse transcriptase yields in the testing pipeline, the DNA agent and their controls will not be affected, but the RNA agent and its controls will fail together." This milestones demonstrates how sensitivity failures are detected by the test. To simulate poor cDNA synthesis, one cDNA containing CHO RNA and one containing CHO RNA spiked with all AA internal standard and native template RNA for MVM5, BPAV1, BTV1 & PTGV2 were mixed to create 100%, 66%, 10%, and 0% cDNA samples. Each mixture was split into 8 samples to which CHO DNA and native DNA agent BPV5 was added. NGS library were created from each sample, sequenced and analyzed (tables below, expected template indicated in blue boxes vs expected RNA native and control AA levels NEAT to 0%). As required by the milestone, the RNA controls and native template yields responded to lower RT levels while the DNA controls were unaffected.

Milestone 3: Internal Standard Detection

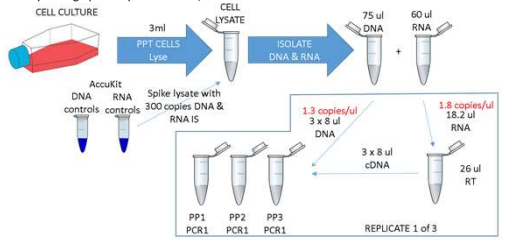


Milestone 3: Native Template Detection



Spike in Considerations

The goal of using the AccuKit reagents was to demonstrate at least 100 copy per ml sensitivity. In the example testing scenario below, 300 IS copies are spiked into lysate created from 3ml cell culture (IS controls should not be spiked into cell culture because nucleases would destroy the controls). The sensitivity controls will be positive if the nucleic acid yield is at least 36% for RNA and DNA. To calculate this minimal purification yields, working backwards, the PCR1 step require ≈10 copies per primer pool, or at least 1.3 copies/ul in the 8 ul DNA added to PCR1 and 1.8 copies/ul in the 18.2 ul RNA added to the RT. This requires purification yields of 98 copies in 75 ul DNA and 108 copies in 60 ul RNA, or at least a 36% purified nucleic acid yield to achieve 100 copy/ml sensitivity. Lower yields may be addressed by scaling up cell input levels and/or the PCR1 volumes.



NEXT STEPS…

- **NCSU will independently validate the AccuKit performance using controls and native templates spiked into CHO lysates.**
- **Bring additional partners to table to provide input into a modular based NGS reagents to meet any regulatory relevant testing for any biomanufacturing platform.**

AccuGenomics Inc. 1410 Commonwealth Dr. Suite 105, Wilmington, NC 28405 www.accugenomics.com