Novel Approaches to ID Testing Using NGS-Based Metagenomics

Robert Schlaberg, MD, MPH

Disclosures

Commercial Interest	What was Received	For What Role
Roche Diagnostics	Honorarium	Advisor
Roche Diagnostics	Research Grants	PI
Hologic	Contract Research	PI
Hologic	Honorarium	Advisor
Epoch Biosciences	Contract Research	PI
Sanofi Pasteur	Contract Research	Co-PI
IDbyDNA	Stock, Honorarium	Co-Founder, CMO, Board Member

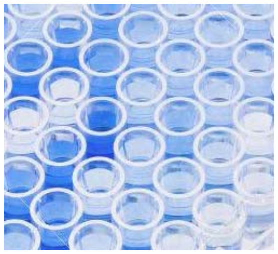
Conventional Infectious Disease Testing

Multimodal

Culture

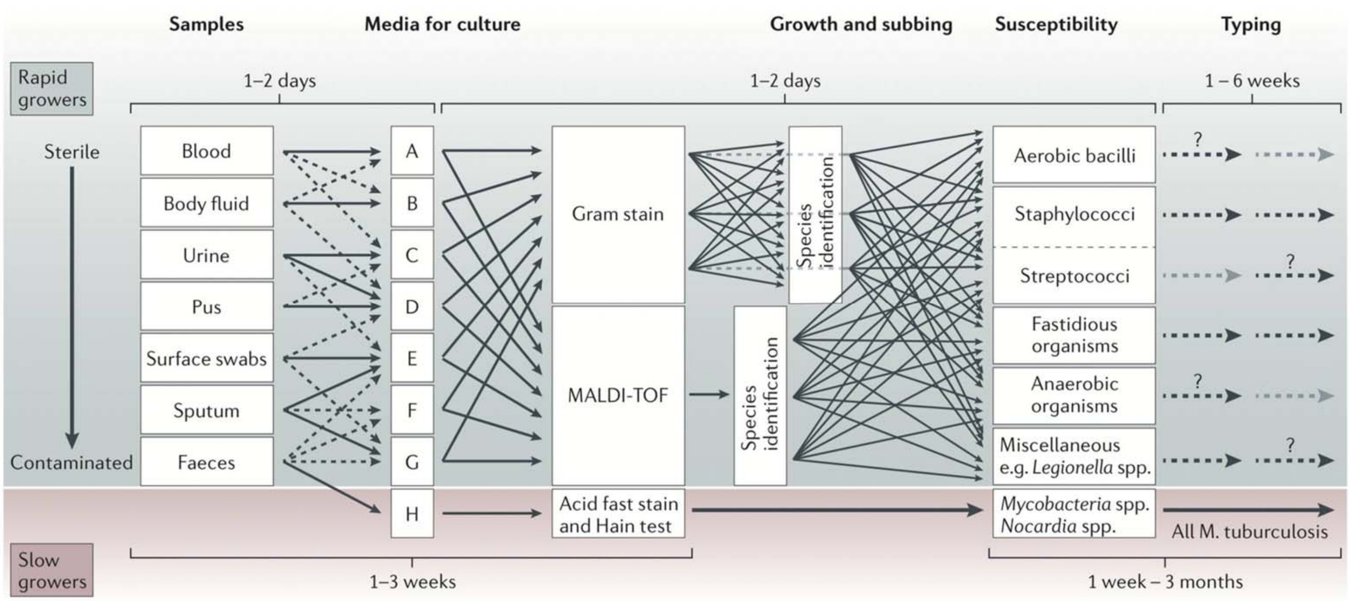
PCR

Ag Detection
Serology



Conventional Infectious Disease Testing

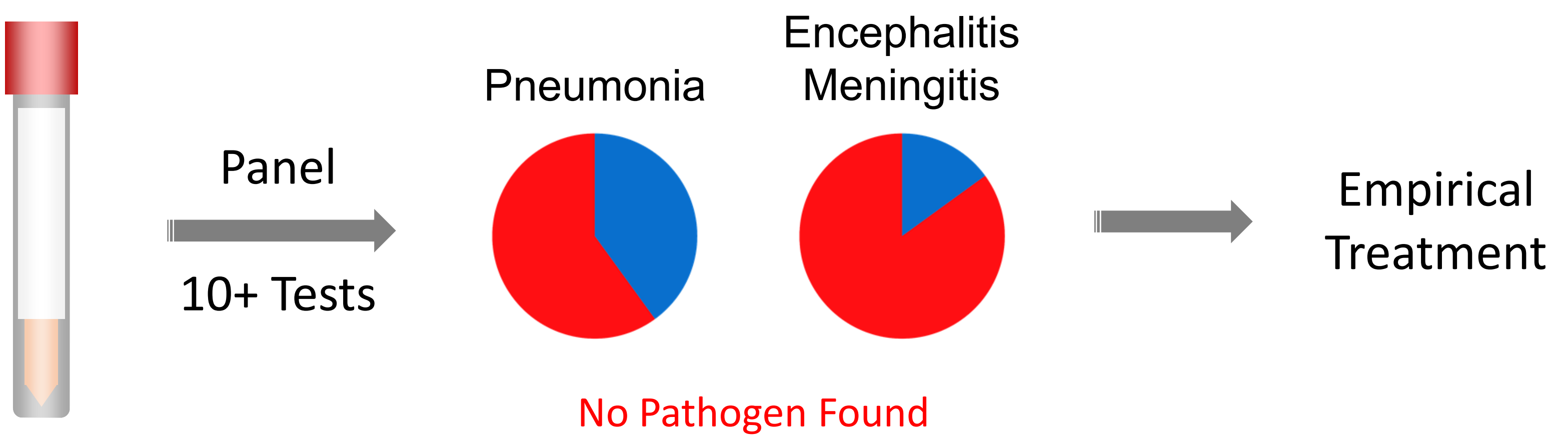
Complex & Sequential Laboratory Workflows



Nat Rev Genet. 2012 Sep;13(9):601-12

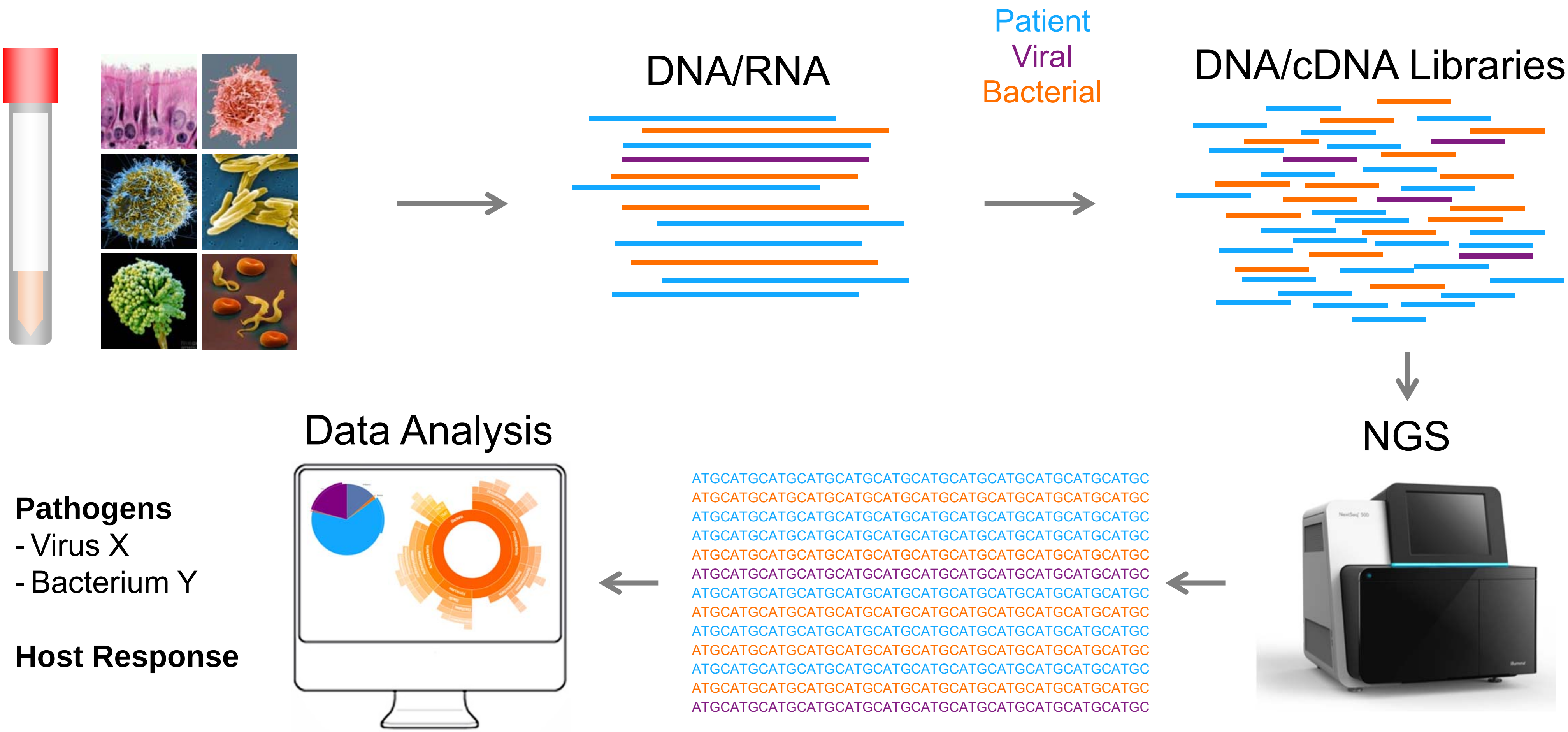
Limited Yield for Many Infectious Syndromes

Current tests often pathogen-specific, combined into panels



Glaser, CID;43(12):1565
Freifeld, CID;52(4):e56
Murdoch, Arch Intern Med.;169(5):463
Jain, NEJM;373(5):415

Metagenomics – One-Stop Testing

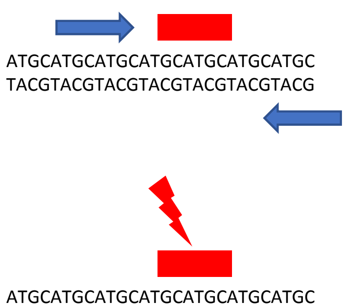


Pathogens
- Virus X
- Bacterium Y

Host Response

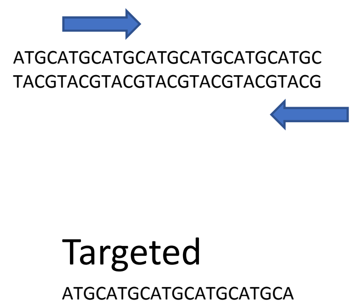
NAAT, Targeted Sequencing, Unbiased Metagenomics

NAAT



- Sensitive
- Specific
- Fast
- Can only detect targeted pathogens

Targeted Sequencing



- Broader scope
- Hypothesis driven
- Classification limited due to genetic snapshot

Unbiased Metagenomics

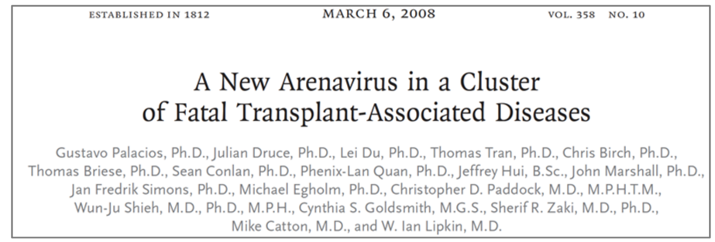
Unbiased

ATGCATGCATGCATGCATGCATGCATGC
TGCATGCATGCATGCATGCATGCATGCA
GCATGCATGCATGCATGCATGCATGCAT
CATGCATGCATGCATGCATGCATGCATG
ATGCATGCATGCATGCATGCATGCATGC
TGCATGCATGCATGCATGCATGCATGCA
GCATGCATGCATGCATGCATGCATGCAT
CATGCATGCATGCATGCATGCATGCATG

- Broadest method
- Hypothesis-free
- Unlimited multiplexing
- All classes of pathogens

Why Do I Think This Is Worthwhile?

First Applications: Pathogen Discovery


ESTABLISHED IN 1812 MARCH 6, 2008 VOL. 358 NO. 10

A New Arenavirus in a Cluster of Fatal Transplant-Associated Diseases

Gustavo Palacios, Ph.D., Julian Druce, Ph.D., Lei Du, Ph.D., Thomas Tran, Ph.D., Chris Birch, Ph.D., Thomas Briese, Ph.D., Sean Conlan, Ph.D., Phenix-Lan Quan, Ph.D., Jeffrey Hui, B.Sc., John Marshall, Ph.D., Jan Fredrik Simons, Ph.D., Michael Egholm, Ph.D., Christopher D. Paddock, M.D., M.P.H.T.M., Wun-Ju Shieh, M.D., Ph.D., M.P.H., Cynthia S. Goldsmith, M.G.S., Sherif R. Zaki, M.D., Ph.D., Mike Catton, M.D., and W. Ian Lipkin, M.D.

Three SOT recipients (same donor)
- Fever, encephalopathy
- Died 4 to 6 weeks after transplantation

Donor
- Cerebral hemorrhage 10 d after travel to rural areas of Eastern Europe

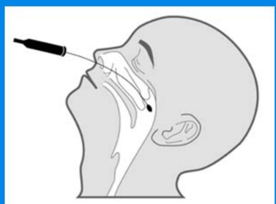
All tests negative
- Bacterial and viral cultures, pathogen-specific PCR's
- Viral and panmicrobial oligonucleotide microarray

Detected by metagenomic cDNA sequencing
- 454 sequencing (~100,000 total sequences)
- LCMV-like arenavirus (protein level)

Palacios et al., NEJM. 2008 Mar 6;358(10):991-8

From Case Reports
to Analytical Performance and Diagnostic Yield

NP Swabs (*n*=109)



PCR Panel (RVP)

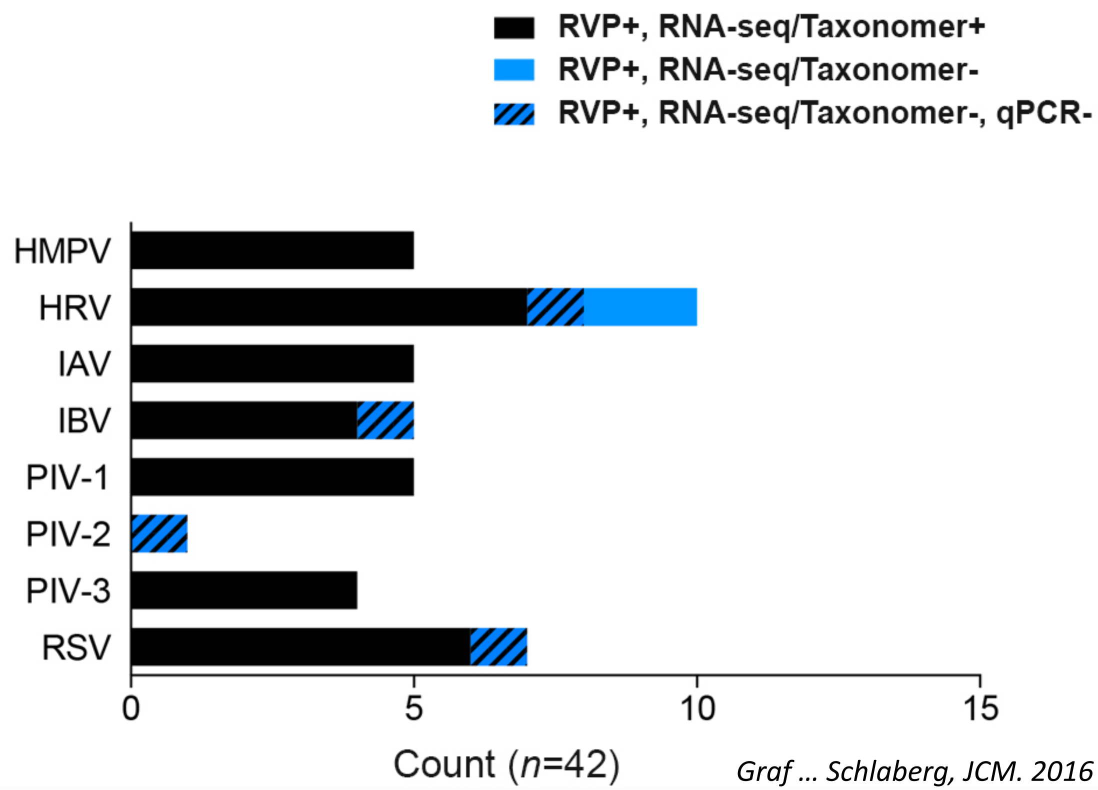
I. Positives (*n*=42)
II. Unselected (*n*=67)

RNA-seq (HiSeq)

5-10 M reads/specimen

>95% Positive Agreement with PCR Panel

NP/OP Swabs (Viral Pathogens), RNA-seq & Taxonomer



Graf … Schlaberg, JCM. 2016 Apr;54(4):1000-7

NP Swabs (*n*=109)

PCR Panel (RVP)

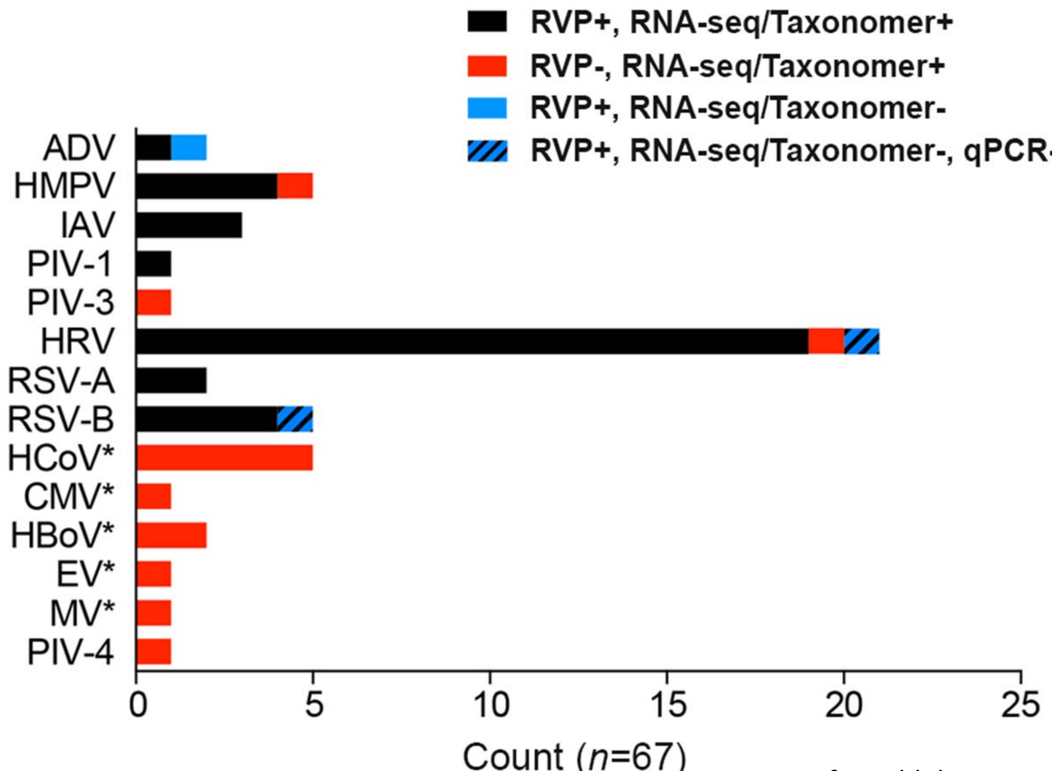
I. Positives (*n*=42)
II. Unselected (*n*=67)

RNA-seq (HiSeq)

5-10 M reads/specimen

>90% Overall Agreement with PCR Panel

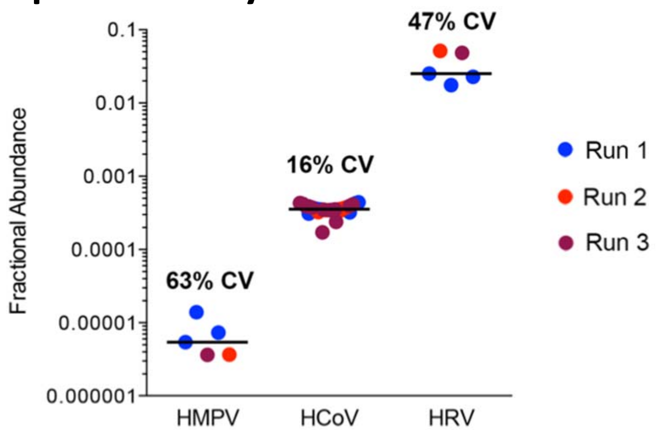
NP/OP Swabs (Viral Pathogens), RNA-seq & Taxonomer



Graf … Schlaberg, JCM. 2016 Apr;54(4):1000-7

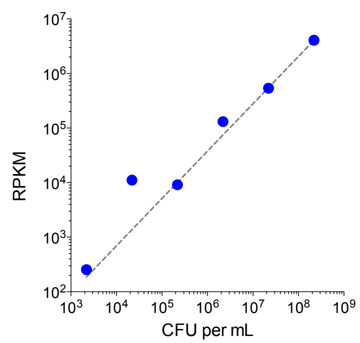
Reproducible & Quantitative Results

Reproducibility

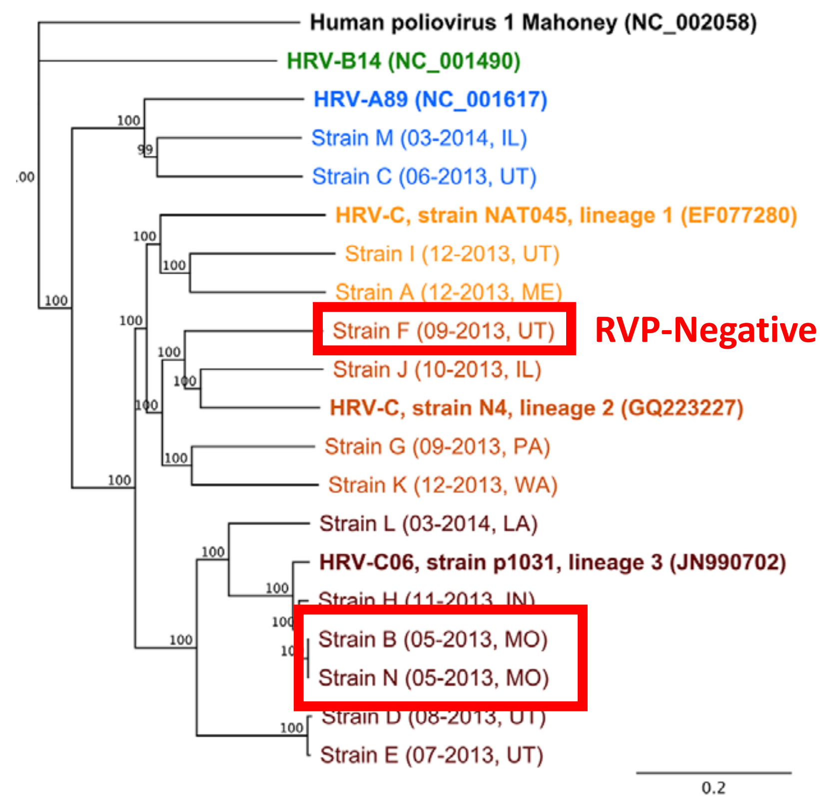


Bacterial Load

H. influenzae



Sequence Agnostic



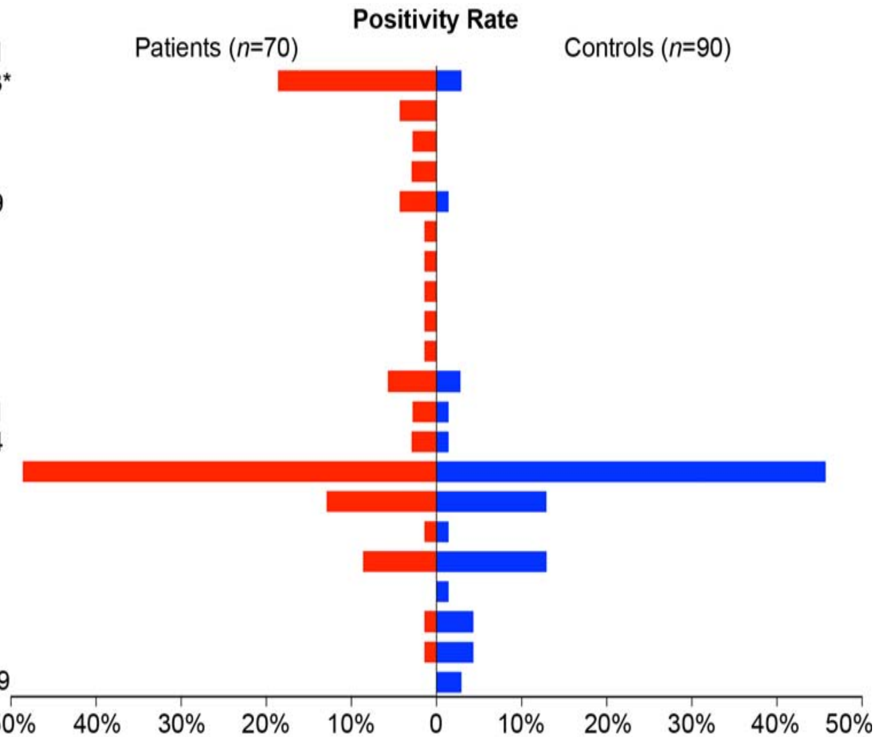
Graf … Schlaberg, JCM. 2016 Apr;54(4):1000-7

CAP of Unknown Etiology: Putative Pathogens in 30%

NP/OP Swabs (Viral Pathogens), RNA-seq & Taxonomer and Panviral Group PCR

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

	OR	95% CI	aOR	95% CI
HBoV	10.0	2.2-46	9.1	1.6-103*
Coxsackievirus	9.4	0.5-185	5.1	0.5-∞
Measles virus	6.6	0.3-140	1.0	0.1-∞
Polyomavirus	6.6	0.3-140	1.3	0.1-∞
HRV-A	4.0	0.4-39	3.5	0.2-199
ADV	3.9	0.2-97	0.1	0.0-∞
C. trachomatis	3.9	0.2-97	0.5	0.0-∞
β-papillomavirus	3.9	0.2-97	1.4	0.0-∞
HSV	3.9	0.2-97	4.0	0.1-∞
Rotavirus	3.9	0.2-97	1.0	0.0-∞
EBV	2.7	0.5-15	2.2	0.2-32
HPIV-4	2.6	0.2-29	9.3	0.4-741
HRV-C	2.6	0.2-29	2.9	0.2-174
Anellovirus	1.7	0.9-3.2	1.6	0.7-3.5
HHV6	1.3	0.5-3.5	1.0	0.3-3.3
Parvovirus B19	1.3	0.1-21	1.0	0.0-88
HHV7	0.8	0.3-2.5	1.2	0.3-4.4
Echovirus	0.4	0.02-11	0.9	0.0-35
CMV	0.4	0.04-4.1	0.5	0.0-8.4
HPeV	0.4	0.04-4.1	0.7	0.0-9.7
Cardiovirus	0.3	0.01-5.1	0.8	0.0-10.9

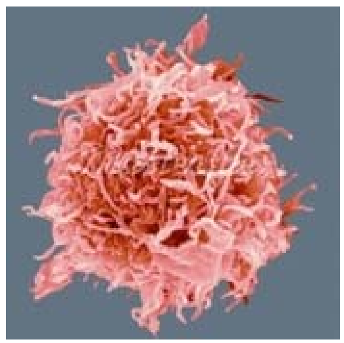


Schlaberg et al., JID;215(9):1407-1415

Can This Be Implemented In Routine Practice?

Specimen-Dependent Performance (Depletion, Enrichment)

Abundant, Unwanted Host Nucleic Acid

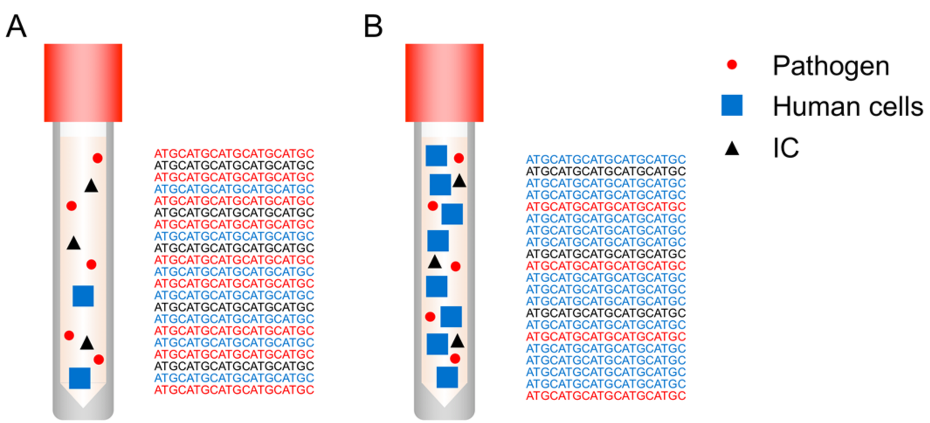


>>

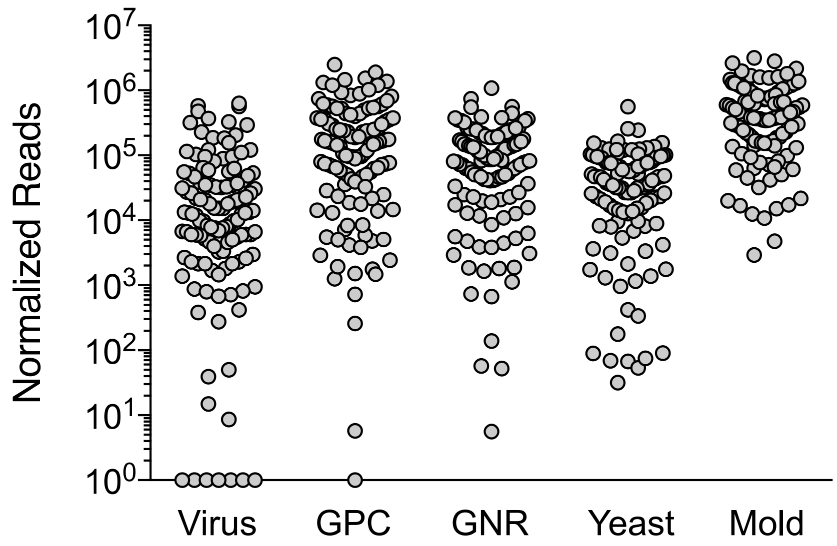
Specimen-Dependent Performance (Depletion, Enrichment)

Analytical sensitivity is variable

- Specimen composition/cellularity
- Sequencing depth
- Genome size



Internal controls to monitor

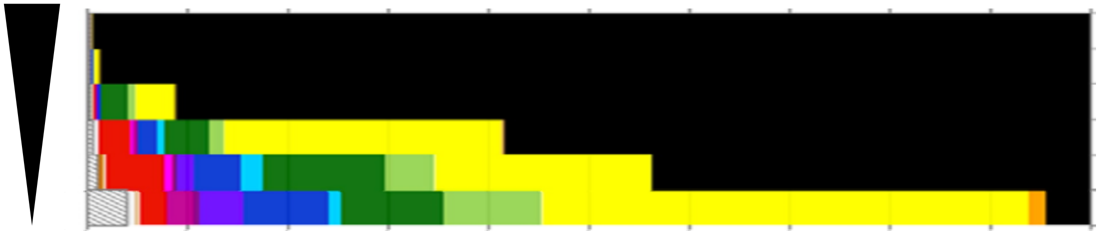


Schlaberg R, Chiu CY, Miller S, Procop G, Weinstock G, Arch Pathol Lab Med. 2017 Jun;141(6):776-786

Contamination Risk (Bacteria), Need for Controls

Reagents contain microbial nucleic acid

- Low DNA Yield -> High Contamination Risk

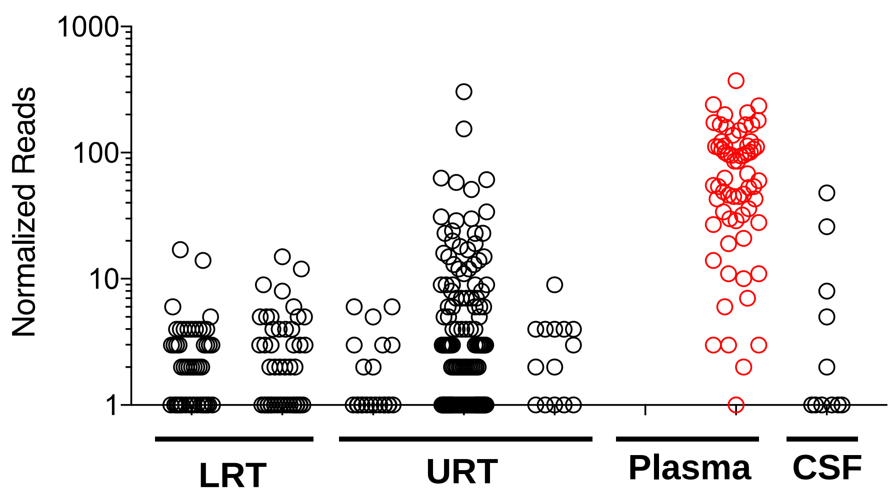


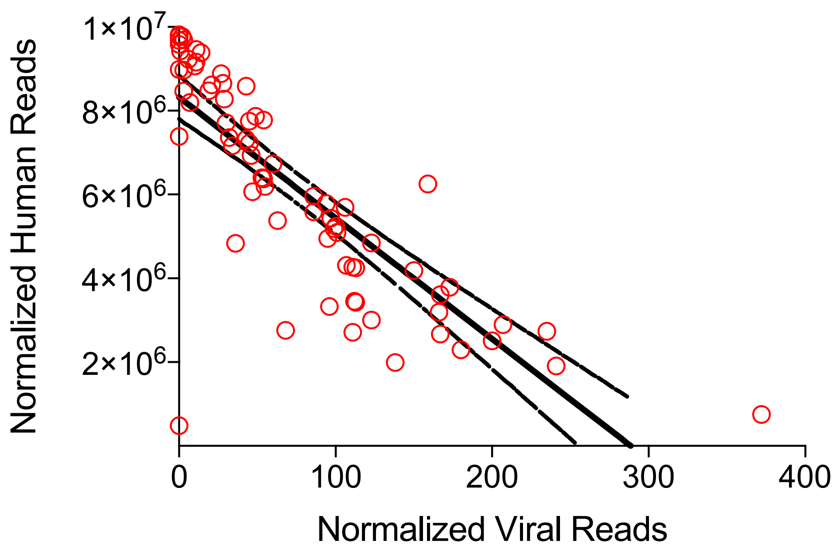
- Highly variable between lots/batches (many reagents, frequent lot changes)

Salter ... Walker, BMC Biol.;12:87
Naccache S ... Chiu CY, J Virol.;87(22):11966-77
Thoendel   Patel, J Clin Microbiol. 2017 Mar 29

Contamination Risk (Viruses), Need for Controls

Kit & NA-Yield Dependent Detection





Salter ... Walker, BMC Biol.;12:87
Naccache S ... Chiu CY, J Virol.;87(22):11966-77
Thoendel Patel, J Clin Microbiol. 2017 Mar 29

Laboratory Workflow Complex to QC

Guaranteeing Consistent Performance

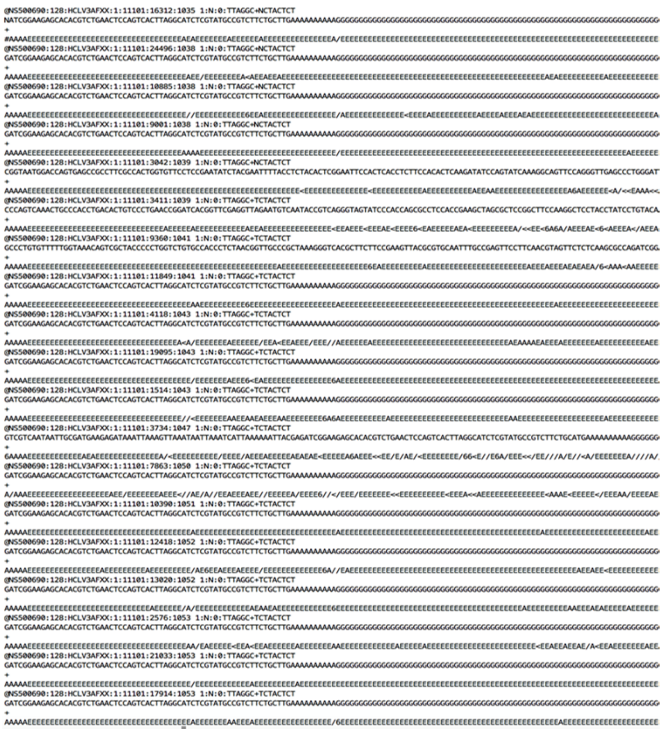
- Involves dozens of steps
- Several steps involve >1 kits/reagents
- Some kits involve reagents from different lots

Rapidly-evolving field

- New, improved reagents & protocols
- Performance comparisons can be complex and expensive
- Update testing protocols frequently

Rapid, Accurate Classification Needed

Millions of DNA Sequences



Millions of Microbial References

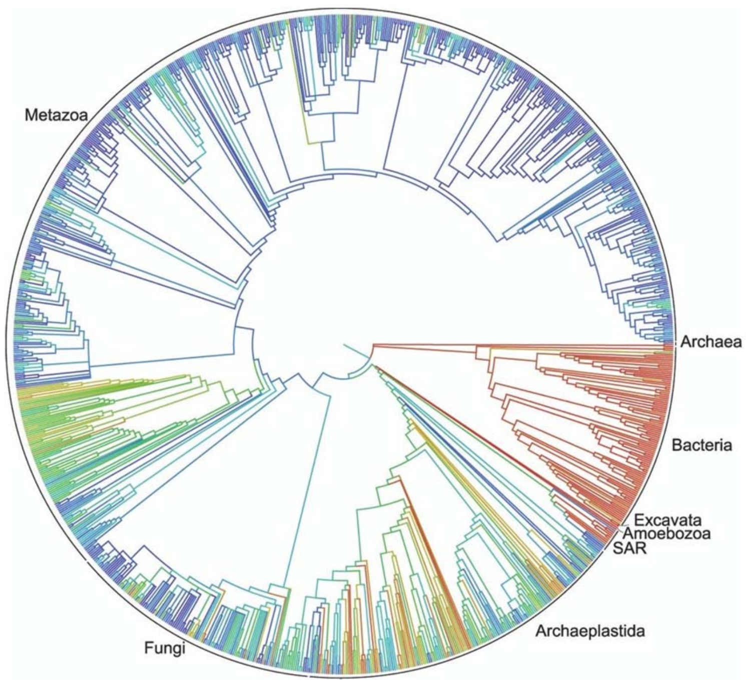
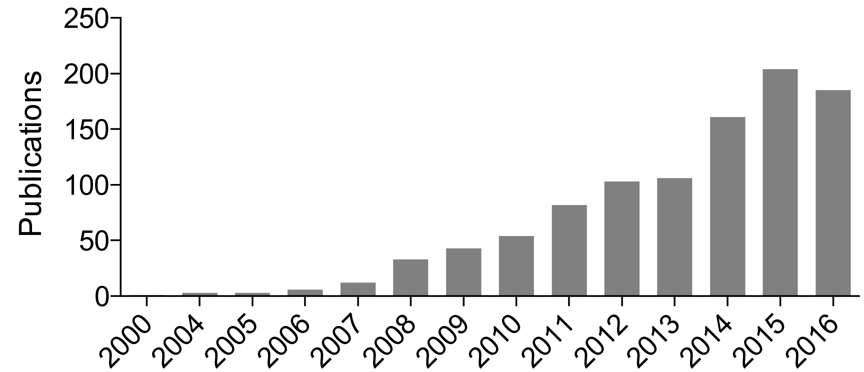


Image from: PNAS;112(41):12764-9

User-friendly Data Analysis Solutions Needed

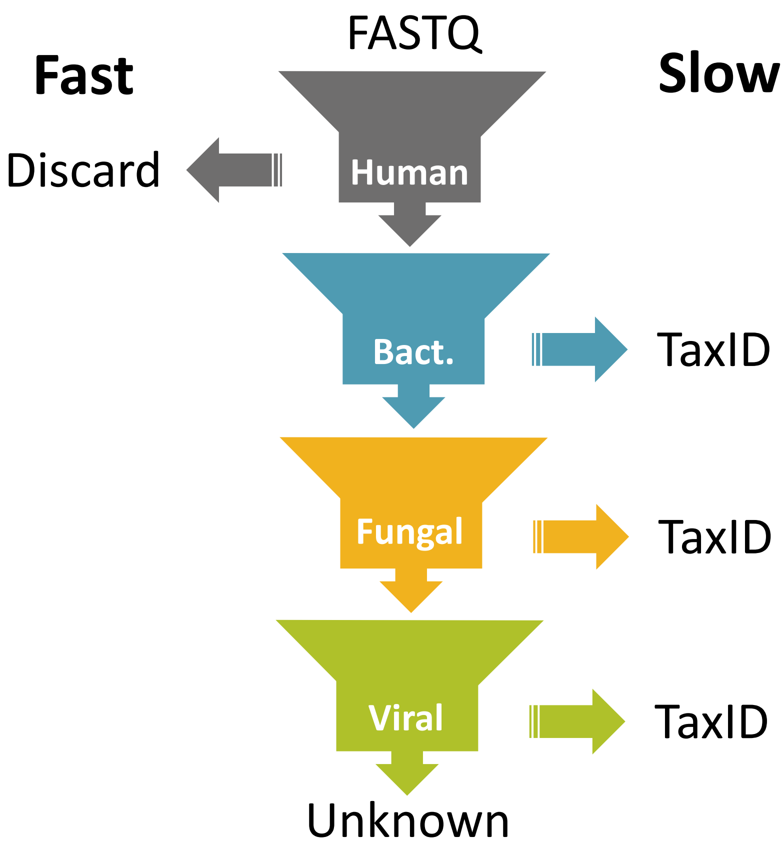
Diagnostic data analysis

- Rapid (minimal impact on TAT)
- Easy to use
- Consistent performance
- Validated
- Easily updatable and expandable



Kraken (Wood, Salzberg; Genome Biol 3;15(3):R46)
SURPI (Naccache, …, Chiu; Genome Res;24(7):1180-92)
CLARK (Ounit, …, Leonardi; BMC Genomics;16:236)
Taxonomer (Flygare, ..., Schlaberg; Genome Biol 26;17(1):111)

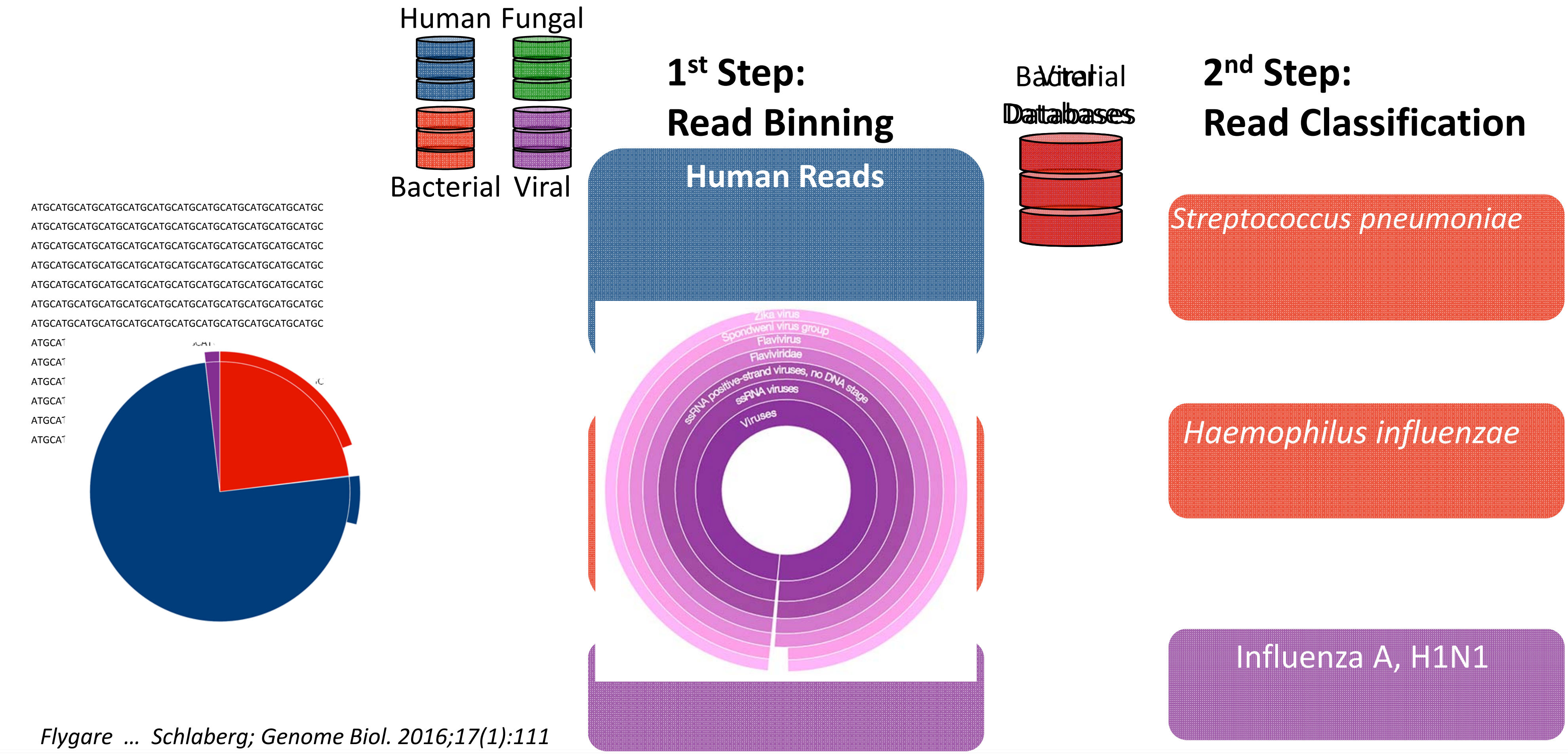
Moving Away from Sequential Subtraction



Disadvantages

- Speed depends on sample composition
- Misclassification
  – Comparison to database subsets
  – Heterogeneous query sequences
  – Difficult to define thresholds

Rapid, Interactive DNA Search Engine (Taxonomer)



Flygare … Schlaberg; Genome Biol. 2016;17(1):111

End-to-End Metagenomics Data Analysis (Explify)

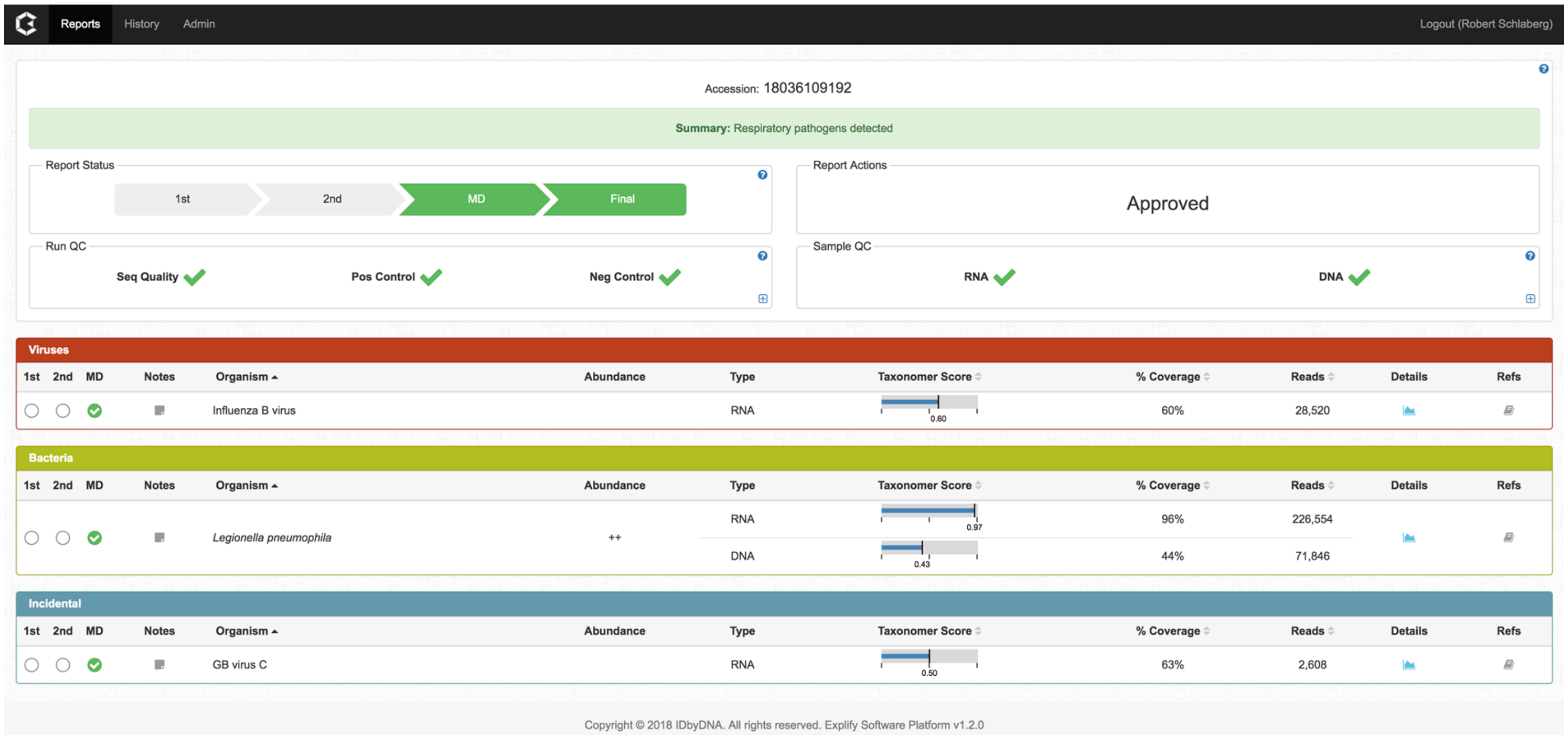
User

- Sequencing Data
- Demultiplexing
- Adapter Trimming, Quality Filtering
- Read Binning, Classification (kmer/Alignment)
- Validated Cutoffs and Interpretive Rules
- Results Visualization in Interactive UI
- Email Alert for New Results

User

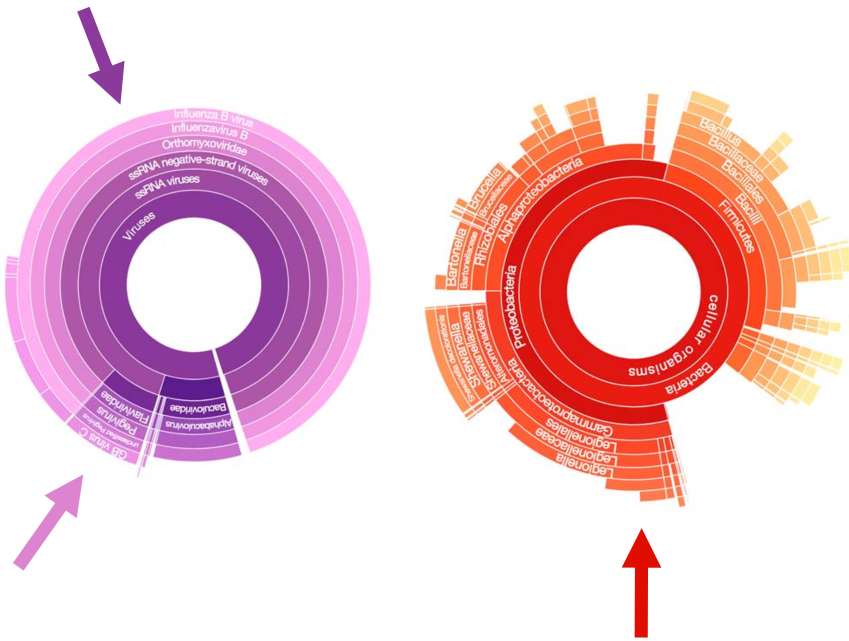
- Result Review and Interpretation
- Documentation, Report Generation

End-to-End Metagenomics Data Analysis (Explify)

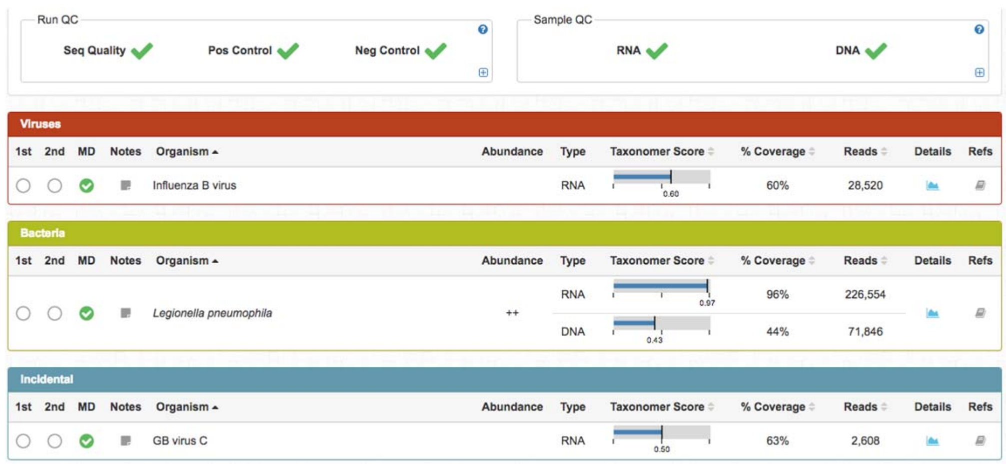


Research vs. Diagnostic Software

Research



Diagnostic



Run QC: Seq Quality ✓ Pos Control ✓ Neg Control ✓

Sample QC: RNA ✓ DNA ✓

Viruses

1st	2nd	MD	Notes	Organism	Abundance	Type	Taxonomer Score	% Coverage	Reads	Details	Refs
		✓		Influenza B virus		RNA	0.60	60%	28,520		

Bacteria

1st	2nd	MD	Notes	Organism	Abundance	Type	Taxonomer Score	% Coverage	Reads	Details	Refs
		✓		*Legionella pneumophila*	++	RNA	0.97	96%	226,554		
						DNA	0.43	44%	71,846		

Incidental

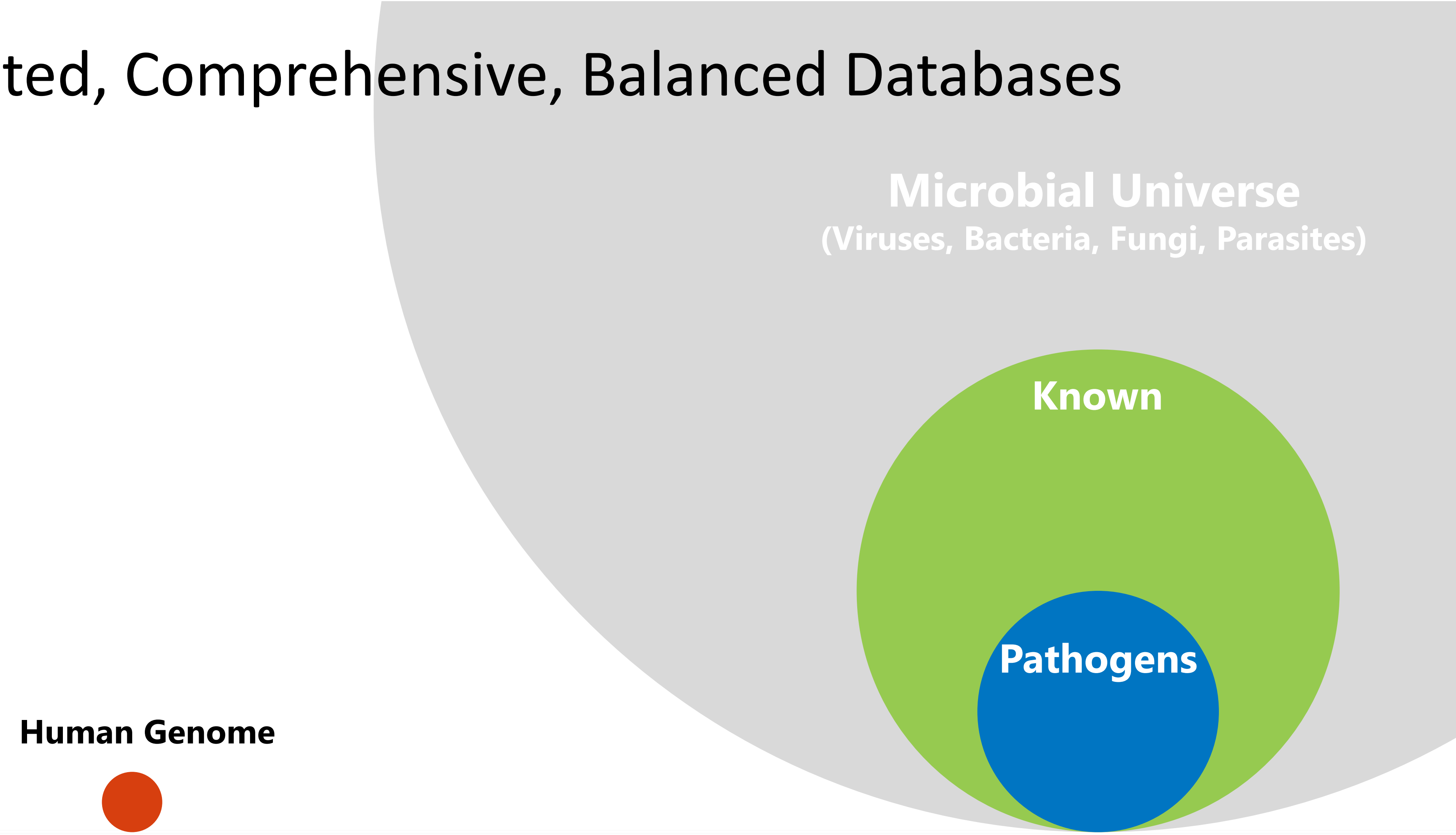
1st	2nd	MD	Notes	Organism	Abundance	Type	Taxonomer Score	% Coverage	Reads	Details	Refs
		✓		GB virus C		RNA	0.50	63%	2,608		

Curated, Comprehensive, Balanced Databases

- Curated
  - Errors: misclassification
- Comprehensive
  - False-negative results if correct match missing
  - False-positive results (near neighbor) if correct match missing
- Balanced
  - Reduce bias

Periodic updates, re-validation, version control

Curated, Comprehensive, Balanced Databases



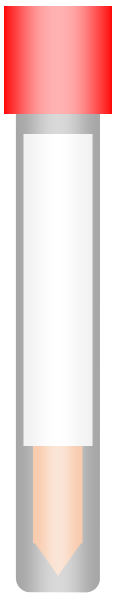
A First Diagnostic Test

Considerations Before Making the Leap

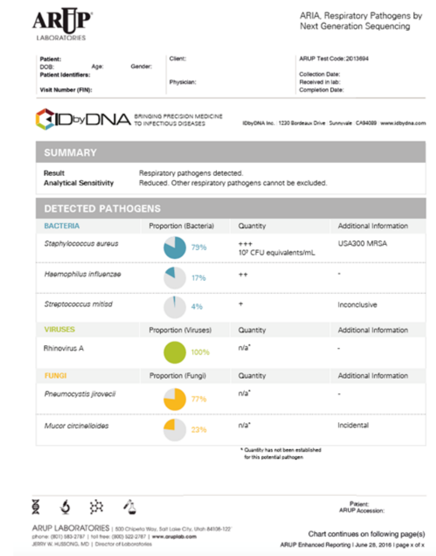
Known Pathogens	Discovery
Standardization	Specimen Variability
Intended Use	Last Resort
Confidence	Promise

Resources (Laboratory, Data Analysis)

First Diagnostic Applications, One Experience



Complete, Accurate, Validated
Relevant TAT
Cost-Effective
Actionable



ARUP LABORATORIES

ARIA, Respiratory Pathogens by Next Generation Sequencing

Patient: DOB: Age: Gender: **Patient Identifiers:** **Visit Number (FIN):**	Client: Physician:	ARUP Test Code: 2013694 Collection Date: Received in lab: Completion Date:

IDbyDNA BRINGING PRECISION MEDICINE TO INFECTIOUS DISEASES — IDbyDNA Inc. | 1230 Bordeaux Drive | Sunnyvale | CA94089 | www.idbydna.com

SUMMARY

Result	Respiratory pathogens detected.
Analytical Sensitivity	Reduced. Other respiratory pathogens cannot be excluded.

DETECTED PATHOGENS

BACTERIA	Proportion (Bacteria)	Quantity	Additional Information
Staphylococcus aureus	79%	+++ 10^7 CFU equivalents/mL	USA300 MRSA
Haemophilus influenzae	17%	++	-
Streptococcus mitisd	4%	+	Inconclusive
VIRUSES	**Proportion (Viruses)**	**Quantity**	**Additional Information**
Rhinovirus A	100%	n/a*	-
FUNGI	**Proportion (Fungi)**	**Quantity**	**Additional Information**
Pneumocystis jirovecii	77%	n/a*	-
Mucor circinelloides	23%	n/a*	Incidental

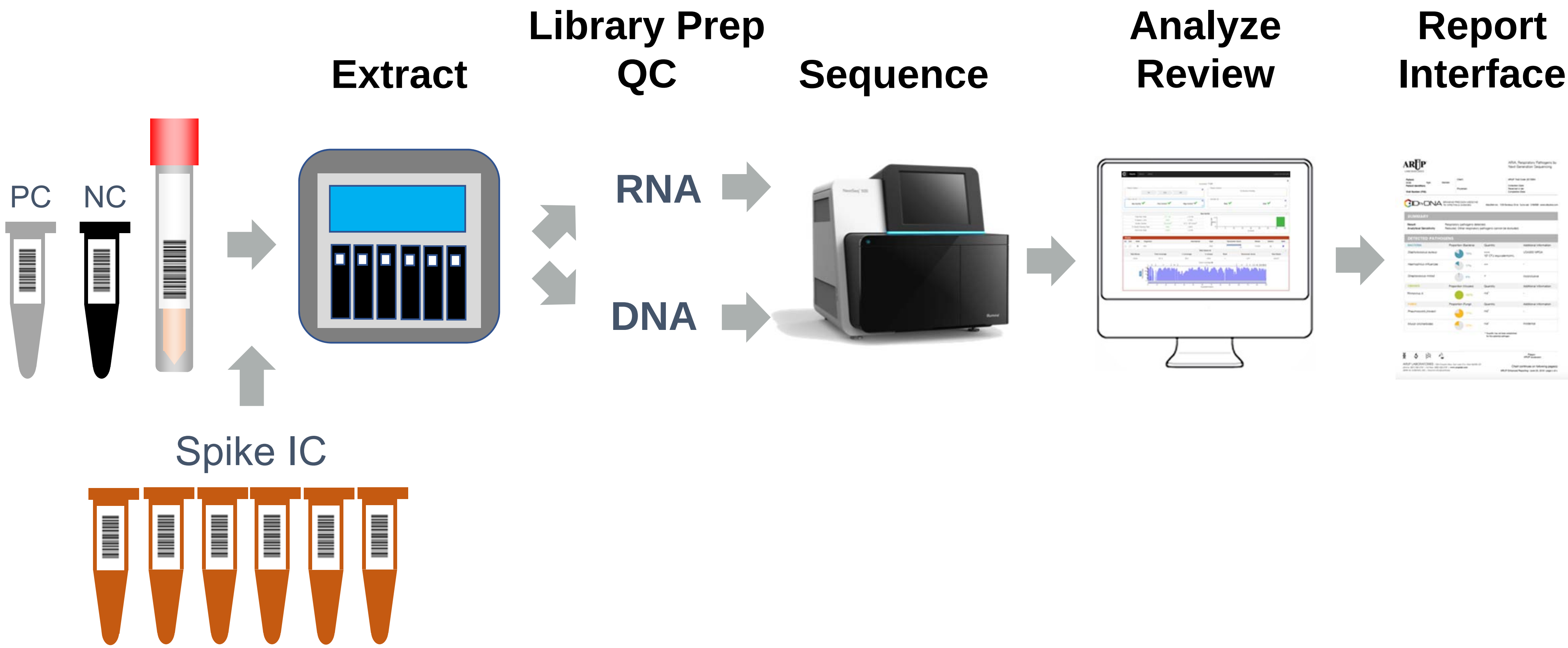
* Quantity has not been established for this potential pathogen

Patient:
ARUP Accession:

ARUP LABORATORIES | 500 Chipeta Way, Salt Lake City, Utah 84108-122
phone: (801) 583-2787 | toll free: (800) 522-2787 | www.aruplab.com
JERRY W. HUSSONG, MD | Director of Laboratories

Chart continues on following page(s)
ARUP Enhanced Reporting | June 28, 2016 | page x of x

Workflow

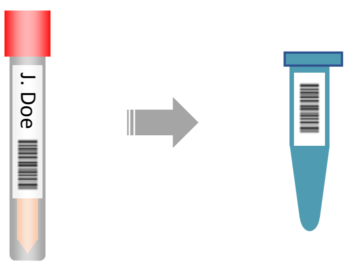

Extract
Library Prep
QC
Sequence
Analyze
Review
Report
Interface
PC
NC
RNA
DNA
Spike IC

Simplifying Diagnostic Workflows

Specimen & Library Prep → **Sequencing** → **Analysis** → **Report**

Validated
- Databases
- Algorithms
- Cutoffs

Wet Bench Workflow



Sequencing

Analysis & Reporting

Validation

200+ validated pathogens

- Viral, bacterial, fungal

Patient Specimens

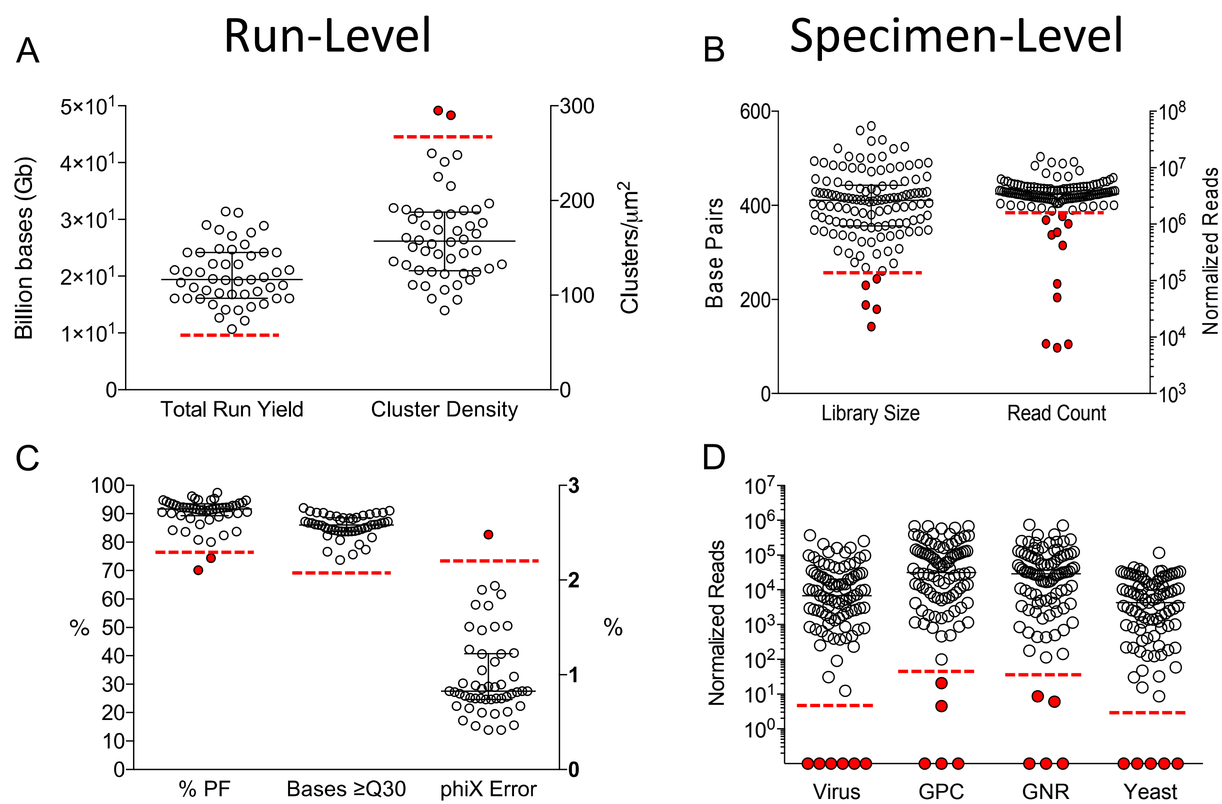
- **Comparison to conventional tests**
- Validation of testing system

1000+ Virtual Specimens

- Pathogens and near-neighbors
- **Databases, algorithms, cutoffs**

QC Metrics

- Run-level
- Specimen-level



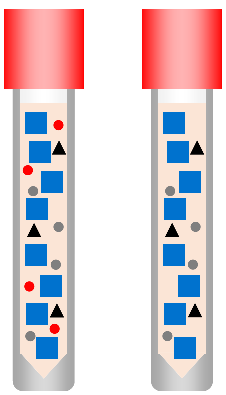
Schlaberg R, Chiu CY, Miller S, Procop G, Weinstock G, Arch Pathol Lab Med. 2017 Jun;141(6):776-786

Validation: Generating Sequencing Reads

- ✓ Sensitivity
- ❑ Specificity
- ❑ Accuracy
- ✓ Reproducibility
- ✓ Stability

External Controls
Internal Controls

Specimens
Pos. Neg.



Wet Lab

Sequencing QC
Specimen QC
Repeat Algorithms

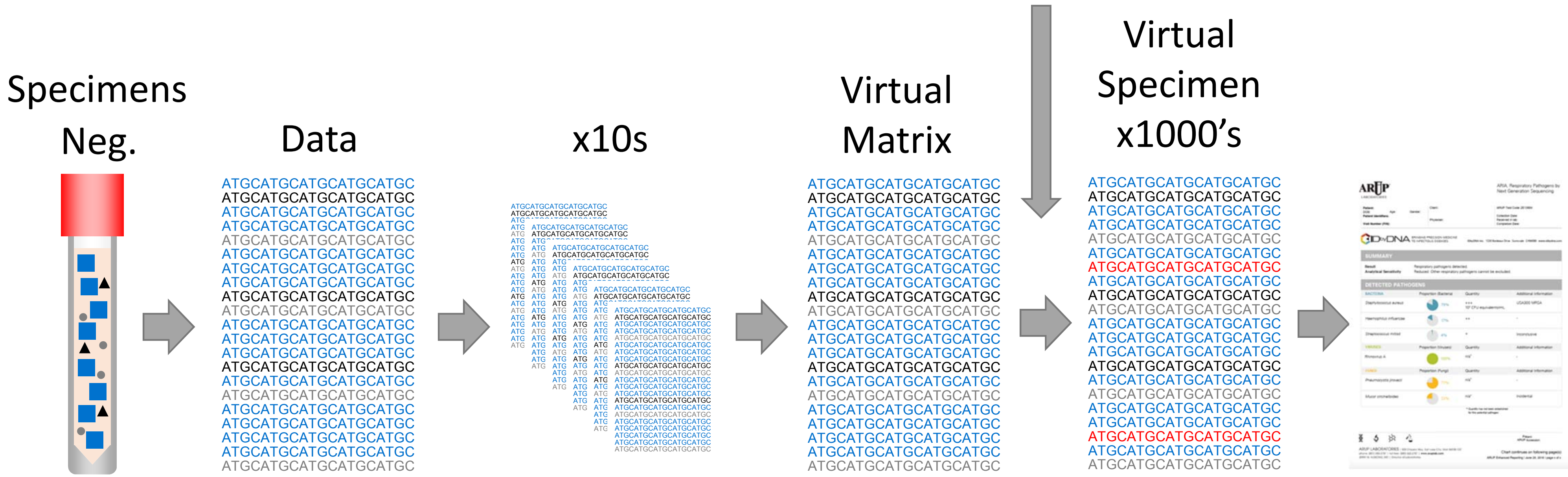
Analysis

Databases
Algorithms
Cutoffs

Validation: Accurate Classification

- ✓ Sensitivity
- ✓ Specificity
- ✓ Accuracy
- ✓ Reproducibility
- ✓ Stability

Pathogens
Near neighbors
Commensals
Contaminants



Diagnostic Yield

Immunocompromised Children with PNA (ICU)

The Children's Hospital of Philadelphia®

Table 1. Underlying Conditions

Condition	n	%
Bone marrow transplantation	17	42
Solid organ transplantation	11	27
Primary immunodeficiency	7	17
Chemotherapy	6	15
High dose steroid therapy	2	5

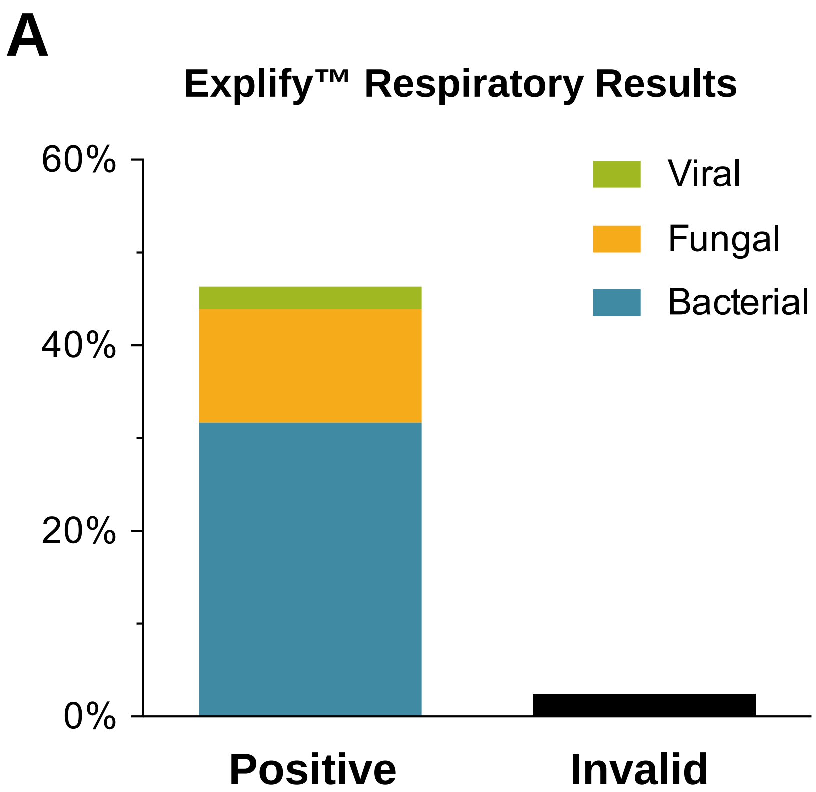
Table 2. Radiographic Findings

Condition	n	%
Diffuse opacities [with consolidation]	28 [7]	68 [17]
Nodules	6	15
Pleural effusions	6	15

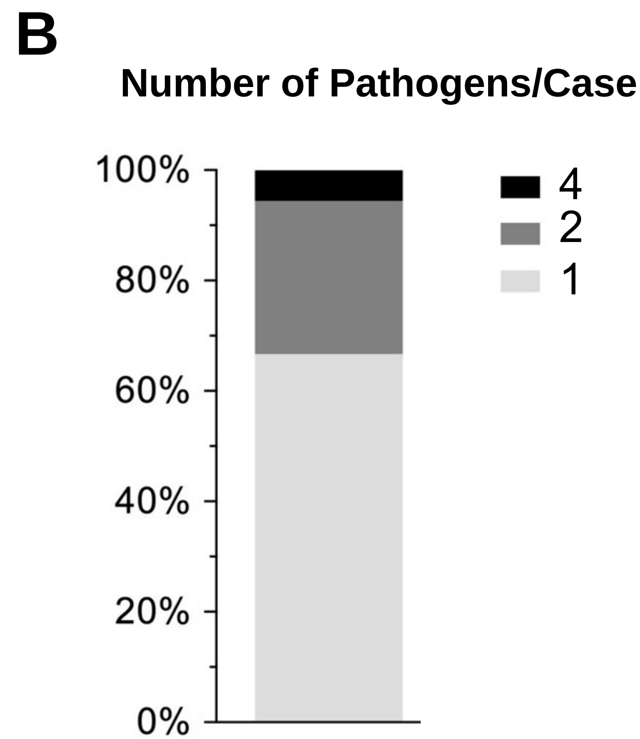
Diagnostic Yield

Immunocompromised Children with PNA (ICU)

The Children's Hospital of Philadelphia



Viruses	Parainfluenza virus Human Bocavirus Rhinovirus
Fungi	*Pneumocystis* *Mucor* *Fusarium*
Bacteria	*Staph. aureus* *Strep. pneumoniae* *S. maltophilia*



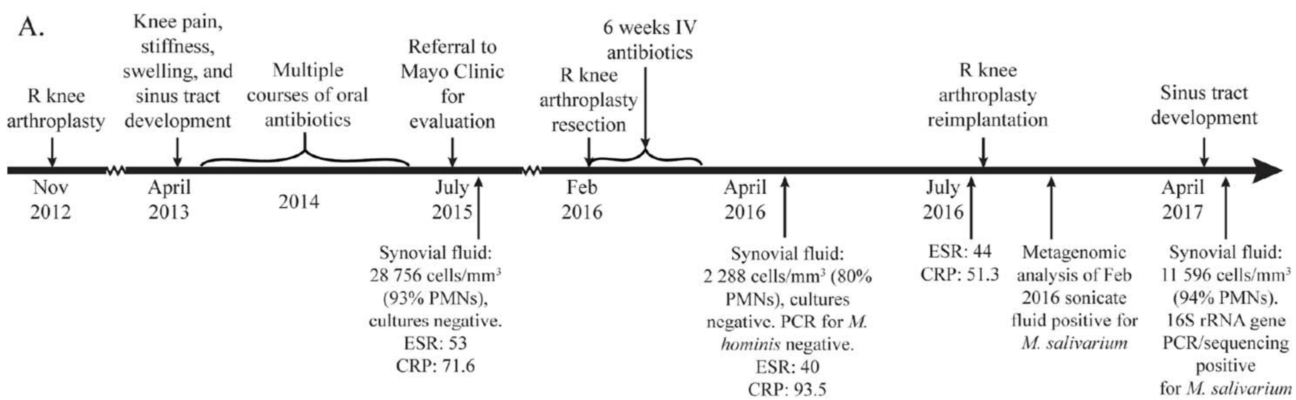
Graf … Schlaberg, manuscript in preparation

Other Applications

Unexpected Pathogens in PJI

52-year-old man with chronic right knee prosthetic joint infection

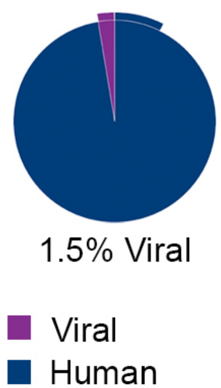
- Swelling, drainage from multiple sinus tracts 6 mo after the TKA
- *Mycoplasma salivarium* by unbiased DNA sequencing



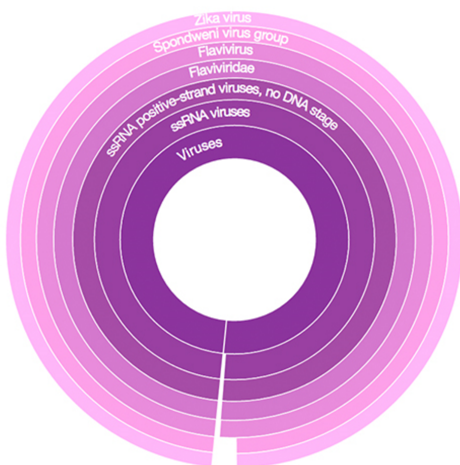
Thoendel, ... , Patel R; Clin Infect Dis. 2017 Jul 15;65(2):332-335

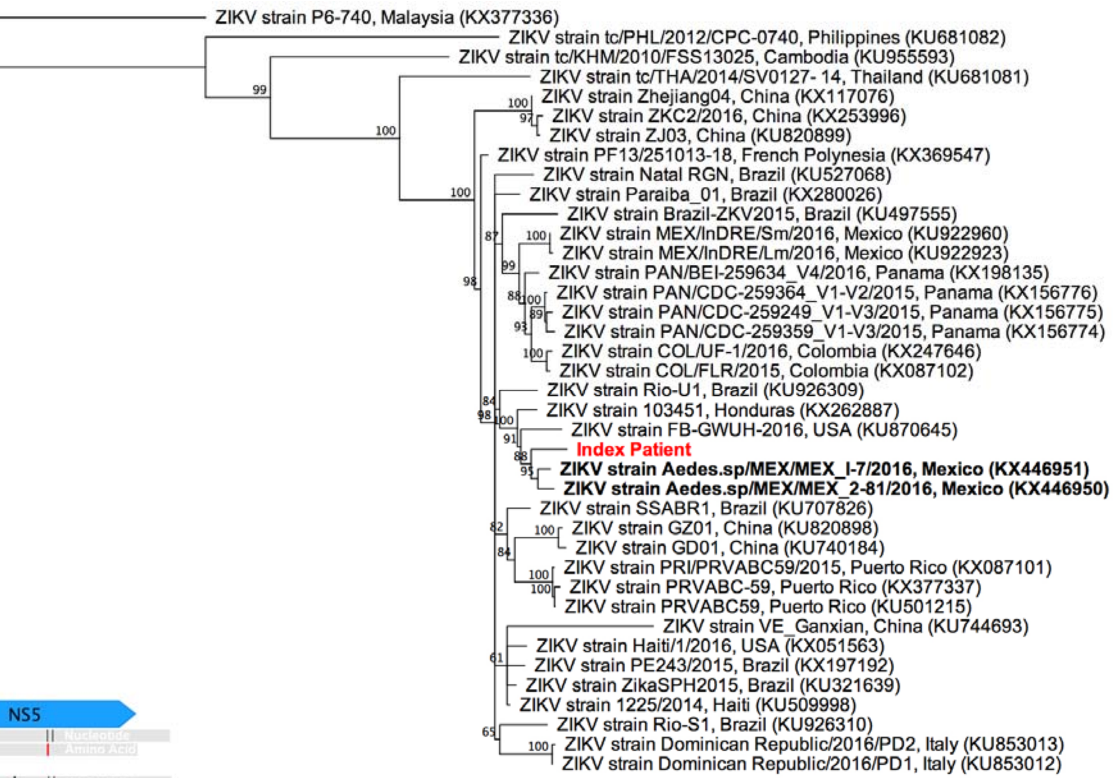
Identification of ZIKV (Only) in Fatal Case

Serum, RNA-seq & Taxonomer



Viral Sequences Identified as ZIKV

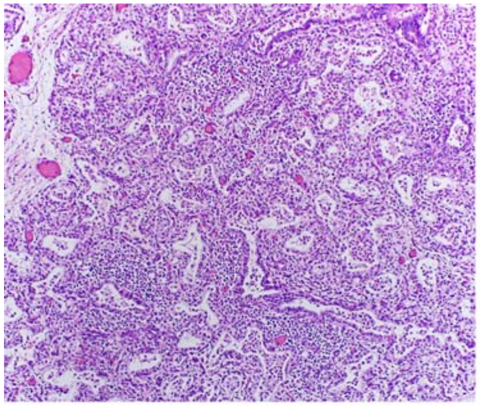




Swaminathan, Schlaberg, et al. September 28, 2016, NEJM

Gestational Psittacosis, Neonatal Death

- 32-year-old pregnant female (16 weeks gestation)
- Persistent fever, neck pain, headache, nausea
- Normal WBC, left shift, thrombocytopenia
- Extensive workup without etiologic diagnosis
- Pre-term labor at 21 weeks gestation
  - Acute chorioamnionitis
  - Congenital pneumonia
- *Chlamydia psittaci* by RNA-seq



Paul, …, Schlaberg, Submitted

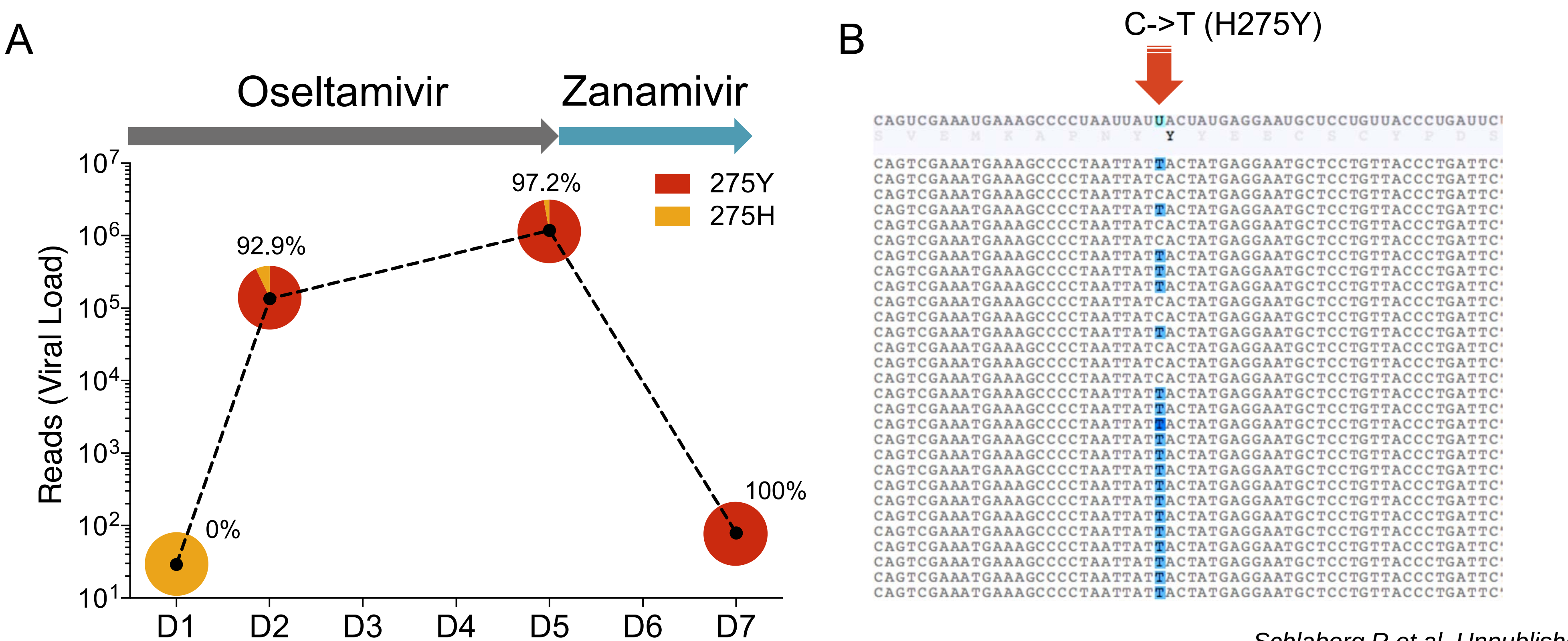
Other Infectious Syndromes

Organ System	Diagnosis	Pathogen (Examples)
CNS	Chronic meningitis	*Leptospira borgpetersenii*
	Progressive encephalitis	Variegated squirrel bornavirus 1
	Meningoencephalitis	*Balamuthia mandrillaris*
Ocular	Chronic uveitis	Rubella virus
	Keratitis	*Fusarium sp.*
Bloodstream	Fever of unknown origin	Parvovirus B19, *P. falciparum*, …
	Sepsis	*Enterobacter cloacae*, …
Gastrointestinal	Diarrhea	Clostridium difficile, …

Simner et al., CID;66(5):778-788
Shigeyasu et al., J Infect Chemother. 2018 Jan 20. pii: S1341-321X(17)30320-3
Grumaz et al., Genome Med. 2016 Jul 1;8(1):73

Getting Closer to A Single-Test Alternative

Monitoring Oseltamivir Resistance, 3-Year-Old Immunocompromised Child



Schlaberg R et al. Unpublished

Much More Information than Pathogen Detection

Presence/Absence of pathogen

Quantification

Drug resistance

Epidemiologic typing

Host-based diagnostics

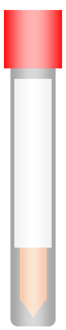
Identification , Quantification, AMR

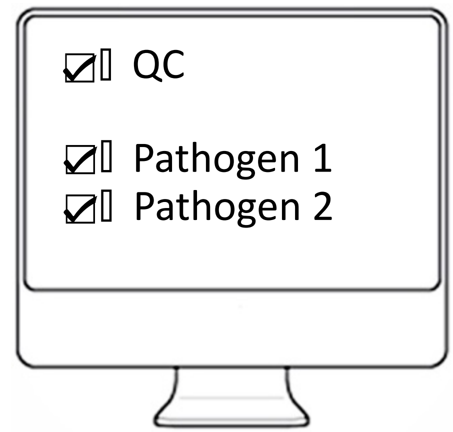
Automated Sample & Library Prep

Rapid Sequencing Flexible Batch Size

Rapid, Automated Data Analysis

Various



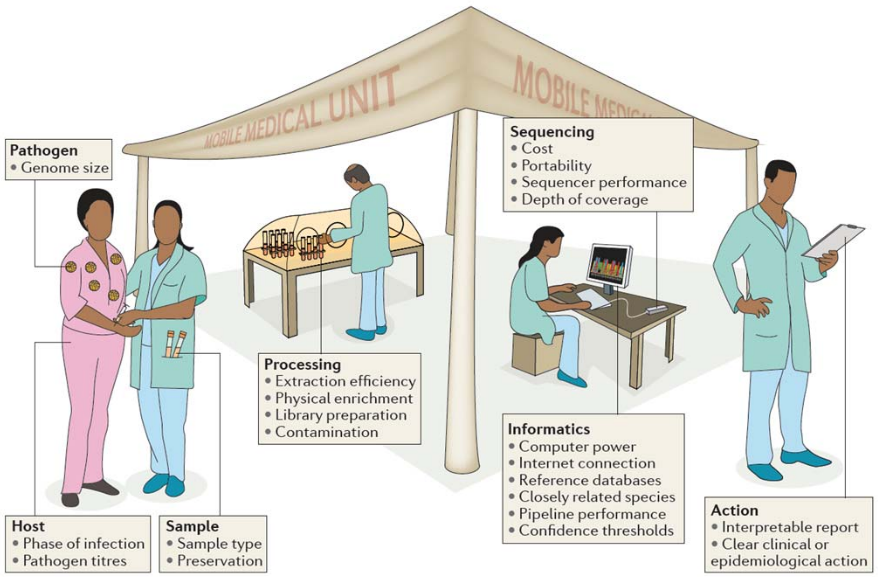


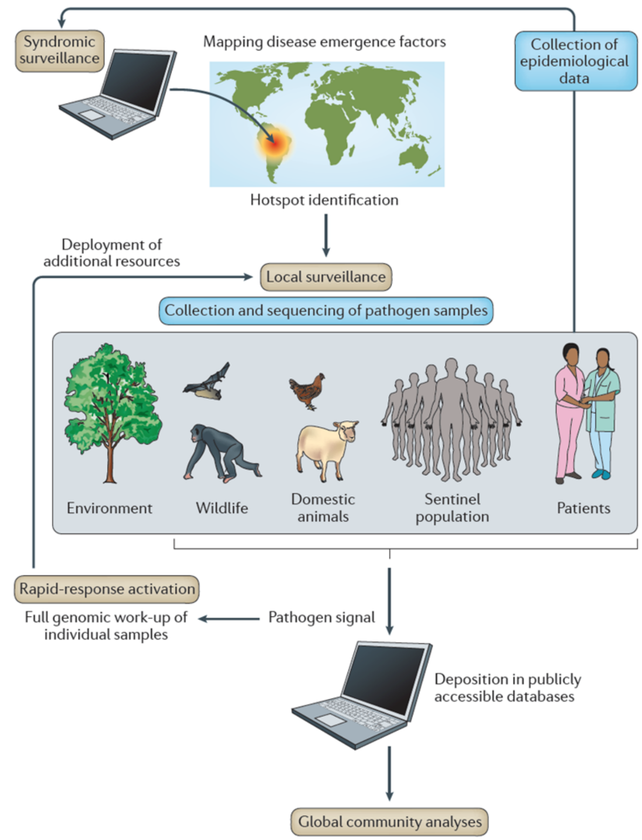
Reagents & Consumables
- DNA/RNA-free
- Closed system

Reagents
- Consistent performance
- Stable reagents
- Few lot changes

Validated
- Databases
- Algorithms
- Cutoffs

Sequencing-based, real-time, global pathogen surveillance





Gardy & Lorman; Nat Rev Genet. 2018 Jan;19(1):9-20

Metagenomics In Routine Diagnostic Practice

Advantages

- Does not rely on differential diagnoses, ‘unbiased’ to get rid of ‘hypothesis-free’
- Resistant to genetic variation, suitable for discovery
- Detection of multiple microorganisms
- Does not rely on viability
- Workflow

Challenges

- Technologies need to become faster, easier to use, cheaper
- Specimen composition affecting performance
- Complex to validate and maintain
- Skills & resources often not available in microbiology laboratories

Allcock et al. J Clin Microbiol. 2017 Nov;55(11):3175-3182

Human Genetics
Brett Kennedy
Mark Yandell

ARUP Laboratories
Keith Tardif
Keith Simmon
Karl Voelkerding
SEQID
Brandy Serrano
Amy Cockerham
Jennie Stanchfield
Diana Mohl

Infectious Diseases
Sankar Swaminathan

Pediatric Infectious Diseases
Chris Stockmann
Carrie Byington
Anne Blaschke
Krow Ampofo
Andy Pavia

Biomedical Informatics
Karen Eilbeck

CHOP
Erin Graf

CDC, Viral Diseases
Seema Jain
Sue Tong
Krista Queen

Pathology
Kim Hanson
Marc Couturier
Maria Pletnikova
Meghan Driscoll

Neurology
Stacey Clardy
Viren Patel

National Institutes of Health
CDC Centers for Disease Control and Prevention
The Children's Hospital of Philadelphia
Bill & Melinda Gates foundation
USTAR Center for Genetic Discovery

Questions?

robert.schlaberg@aruplab.com