# Analytical Validation of Microbial Hybrid Capture Sequencing from Multiple Sample Types

Prepared by: Curtis Bacon, PhD
Eurofins Viracor
May 12, 2025

NeoGen Fungal / AFB NGS

## KEY POINTS

- Hybrid capture is a sensitive and selective method for diagnosing invasive fungal/bacterial infections (IFBI).
- Utilization of robust cutoff criteria effectively filters hybrid capture false-positive results while still maintaining a low false-negative rate.
- Selective targeting of clinically relevant fungi and bacteria allow for simple-to-interpret resulting, even in complex samples such as BAL.

## Introduction

Fungal and bacterial infections can lead to serious complications for immunosuppressed patients[1]. Rapid and accurate identification of the etiological agent involved in invasive fungal/bacterial infections (IFBIs) can allow for reduced time to initiation of the appropriate targeted therapy and preemptively eliminate starting therapies that could ultimately prove ineffective. However, current IFBI diagnostics suffer from low sensitivity/specificity, invasiveness (e.g., biopsy), have slow turnaround times (e.g., culture), and/or require an *a priori* hypothesis regarding which pathogen has infected the patient (e.g., pathogen-specific PCR based methods)[2].

Circulating cell-free DNA (cfDNA) has recently become an attractive analyte due to the non-invasive collection procedure compared to standard methods[3,4]. cfDNA can originate from various sources within the body and through many differing pathways including apoptosis, necrosis, tumor cells, and microorganisms. Extraction of circulating cfDNA from a non-invasive collection of patient blood samples followed by next generation sequencing (NGS), which provides an unbiased assessment of the organisms present in a sample, could overcome the current shortcomings in IFBI diagnostics. However, current metagenomic cfDNA NGS assays lack clinical impact, particularly for patients infected with fungal or non-tuberculosis (NTM) bacteria, largely due to low sensitivity and/or specificity[5,6]. Additionally, diagnosis of the etiological pathogen using metagenomic cfDNA NGS can be difficult when multiple organisms are reported[7].

While a non-invasive plasma cfDNA assay is ideal, clinical situations may call for testing a sample taken directly from the site of infection, such as a bronchoalveolar lavage (BAL). Performing genomic DNA NGS directly on respiratory samples would likely increase the chance of detecting a respiratory pathogen because pathogen burden is expected to be higher in the compartment of origin. However, determining the causative pathogen can be complicated due to the presence of a natural flora and higher likelihood of polymicrobial detection[7].

Targeted enrichment of medically relevant pathogens using a hybrid capture method addresses many current issues plaguing metagenomic NGS methods by facilitating higher sequencing coverage of targeted, clinically relevant organisms[8]. Tens of thousands of hybrid capture probes can be incorporated in a single assay thereby allowing for the selective detection of a targeted yet large breadth of pathogens[8]. Furthermore, multiple probe sets targeting different pathogen panels can be developed and run either in isolation or in conjunction with each other allowing for a customized diagnostic approach in accordance with the presentation of the patient, thereby potentially reducing the likelihood of non-etiological organisms being reported. Described here is a sensitive and specific hybrid capture diagnostic that selectively targets relevant pathogens, producing straightforward results that aid in the diagnosis IFBIs.

## Methods & Applications

### Hybrid Capture Method for Diagnosis of IFBIs

A hybrid capture assay was developed that sequences cell-free DNA (cfDNA) from plasma or genomic DNA (gDNA) from BAL that has been enriched by a set of 25,665 biotinylated hybrid capture probes targeting a panel of 673 fungi and bacteria (Figure 1A). The sequencing data is analyzed using a custom developed bioinformatic pipeline that filters out human aligned reads, assembles reads aligned to a hand-curated pathogen database, and reports organisms that pass a set of cutoff criteria established during validation (Figure 1B).

**Figure 1A**

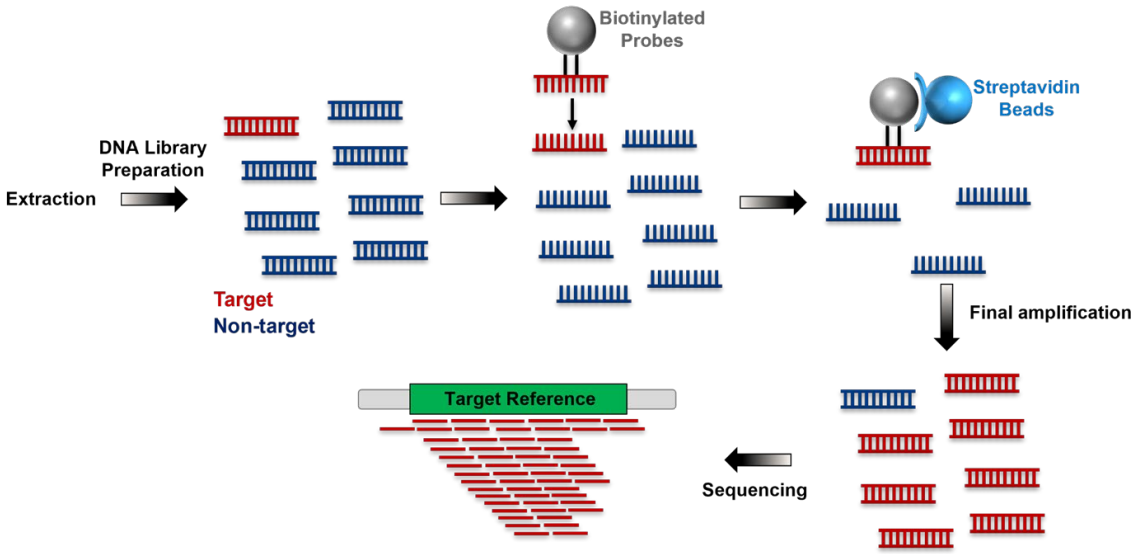


**Figure 1B**

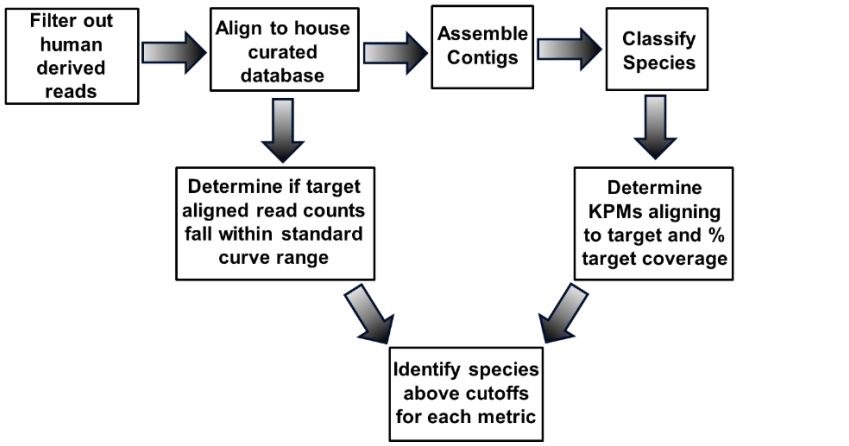


**Figure 1.** ***Overview of the Hybrid Capture Method.*** Overview of the **A)** Bench Workflow and **B)** Bioinformatics Pipeline

## Establishment and Validation of Cutoff Criteria

Robust positive cutoff criteria were developed using limit-of-blank and contrived sample data derived from plasma samples in conjunction with receiver operator curve (ROC) and standard curve analysis. The NexGen™ analysis pipeline outputs three metrics for each detected organism; kmers/million (KPM) derived from pathogen DNA, percent target coverage, and read counts. ROC analysis for KPM and percent target coverage yielded an AUC of 0.89 and 0.92 (95% CI 0.85-0.92 and 0.89 to 0.95), respectively, thereby demonstrating the hybrid capture analysis pipeline's ability to separate true positives from false positives based on KPM and percent target coverage (Figure 2A). Cutoffs for both KPM and percent target coverage were determined by selecting the value that predicted >80% specificity and >90% sensitivity.

**Figure 2A**

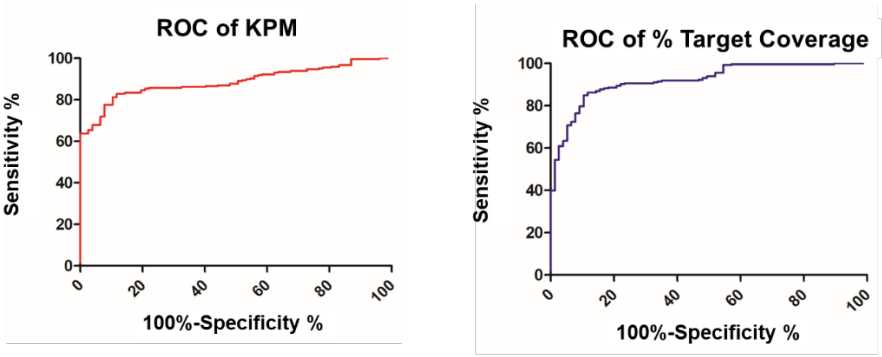


**Figure 2B**

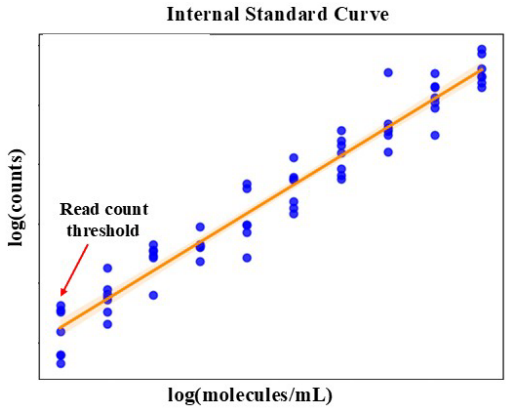


**Figure 2.** ***Establishment of the Cutoff Criteria.*** **A)** ROC analysis comparing negative plasma samples and plasma samples contrived with select organisms of interest to determine the cutoff criteria for KPM and percent target coverage. **B)** Example of an internal standard curve. Each sample is spiked with a pool of internal standards mixed at varying concentrations. For a detected target to pass the counts cutoff, its aligned read counts must surpass the average number of read counts aligning to the lowest detected internal standard.

A third cutoff utilizes read counts aligned to each detected organism and a sample specific internal standard curve. The read count cutoff value is obtained on a sample-by-sample basis by determining if the number of reads aligning to a detected target are greater than or equal to the lowest number of reads that align to the lowest concentration of internal standard present in each sample (Figure 2B).

After defining cutoff criteria for each pipeline metric, analytical validation was performed on plasma samples contrived with fungi and bacteria representative of the >600 species targeted by the hybrid capture probe set (Table 1).

| Organisms | Kingdom | Characteristic |
|---|---|---|
| *Aspergillus fumigatus* | Fungi | Mold |
| *Aspergillus terreus* | Fungi | Mold |
| *Aspergillus niger* | Fungi | Mold |
| *Fusarium solani* | Fungi | Mold |
| *Candida tropicalis* | Fungi | Yeast |
| *Candida kefyr* | Fungi | Yeast |
| *Candida rugosa* | Fungi | Yeast |
| *Candida intermedia* | Fungi | Yeast |
| *Candida krusei* | Fungi | Yeast |
| *Candida pelliculosa* | Fungi | Yeast |
| *Candida lusitaniae* | Fungi | Yeast |
| *Candida parapsilosis* | Fungi | Yeast |
| *Candida auris* | Fungi | Yeast |
| *Candida glabrata* | Fungi | Yeast |
| *Candida albicans* | Fungi | Yeast |
| *Pneumocystis jirovecii* | Fungi | Dimorphic |
| *Cryptococcus neoformans* | Fungi | Dimorphic |
| *Cryptococcus gattii* | Fungi | Dimorphic |
| *Mycobacterium avium* | Bacteria | NTM* |
| *Mycobacterium abscessus* | Bacteria | NTM |
| *Mycobacterium kansasii* | Bacteria | NTM |
| *Nocardia asteroides* | Bacteria | AFB** |
| *Nocardia farcinica* | Bacteria | AFB |
| *Actinomyces israelii* | Bacteria | nAFB† |

* Non-tuberculosis bacterium

** Acid-fast Bacteria

*†Non-acid-fast Bacteria*

**Table 1. *Organisms Validated at the Bench.*** Organisms spanning both the fungi and bacteria kingdoms and with varying characteristics were validated at the bench using the hybrid capture workflow.

In the cfDNA plasma dataset, 80% sensitivity and 98% specificity were observed when applying three-cutoff metrics to the validation data. However, when an identified pathogen was required to meet the criteria of two out of the three cutoff metrics, the sensitivity increased to 98% with only a marginal drop in specificity to 96%. Note that when calculating specificity, a result was considered a 'true negative' if there were no organisms identified in addition to the specific target organism contrived in the sample. If an additional organism was identified, then the sample was counted as a 'false positive'. Similar results (95% sensitivity and 97% specificity) were observed for samples contrived in BAL.

To further validate the developed cutoff criteria for hybrid capture, a clinical accuracy sample set was analyzed by both hybrid capture and in-house qPCR assays. Fourteen residual patient BAL samples were processed using the hybrid capture protocol, twelve of which were positive for an organism that also has a validated in-house qPCR assay available (Table 2). 100% agreement was observed between hybrid capture and qPCR. Notably, the hybrid capture assay was able to detect pathogens that yielded qPCR Ct values as high as 35 (Table 2), further demonstrating the sensitivity of the hybrid capture assay.

| Sample | Organism(s) Detected by NexGen and Tested by qPCR | qPCR Result | qPCR Ct |
|---|---|---|---|
| 1 | *Nocardia farcinica* | Detected | 22.7955 |
| 2 | *Candida albicans; Candida parapsilosis* | Detected; Detected | 30.1461; 20.142 |
| 3 | *Candida tropicalis* | Detected | 23.8048 |
| 4 | *Candida tropicalis* | Detected | 26.346 |
| 5 | *Aspergillus spp.* | Detected | 18.9157 |
| 6 | *Aspergillus spp.* | Detected | 17.9816 |
| 7 | *Aspergillus spp.* | Detected | 29.0888 |
| 8 | *Aspergillus spp.* | Detected | 30.9734 |
| 9 | *Candida albicans; Candida tropicalis* | Detected; Detected | 29.9633; 35.797 |
| 10 | *Candida albicans* | Detected | 26.7835 |
| 11 | *Aspergillus spp.* | Detected | 27.1402 |
| 12 | *Candida albicans* | Detected | 29.1742 |

**Table 2. *Comparison of Hybrid Capture and qPCR Results in Clinical BAL Samples.*** Fourteen residual clinical BAL samples were tested using hybrid capture. Twelve of these samples were positive for an organism that also has an in-house validated companion qPCR assay. All twelve samples were positive by both hybrid capture and qPCR.

## Challenges Associated with the Complexity of BAL

Polymicrobial results complicate a clinician's ability to interpret diagnostic assay reports[7]. While the hybrid capture assay displayed high specificity regarding detection of multiple organisms in plasma, the likelihood of detecting multiple organisms in a BAL sample is higher due to the presence of a natural flora[9,10]. To characterize the number and type of organisms that are expected to be frequently detected in BAL samples, 111 residual patient BAL samples were screened using hybrid capture. Up to two organisms were detected in >60% of the samples, indicating that the targeted nature of the hybrid capture helps to limit the number of polymicrobial reports produced compared to a whole genome sequencing methodology which could result in large numbers of normal oral and respiratory flora (Figure 3A).

Some organisms were frequently observed in the BAL specimens analyzed for this study. Figure 3B lists those that were seen in at least 5% of BAL specimen analyzed. *Actinomyces spp., Candida spp., and Aspergillus spp.* made up most of the organisms detected in at least 5% of the screened BAL samples (Figure 3B). This result is perhaps not surprising considering several species of Actinomyces are known to colonize the respiratory tract[10]. The common *Actinomyces* co-colonizer, *Aggregatibacter actinomycetemcomitans*, was also detected in 16% of the samples[10]. It should be noted that the residual BAL samples used for this study were taken from samples submitted to a reference laboratory for diagnostic testing by other methods, therefore, a portion of the detected species can likely be attributed to a patient's disease state. Nonetheless, this data demonstrates that a relatively small set of potential colonizer organisms are targeted by the hybrid capture assay, however, the organisms in Figure 3B can cause disease in rare conditions and therefore were left in the database.

**Figure 3A**

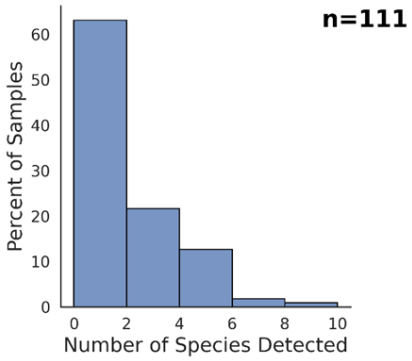


**Figure 3B**

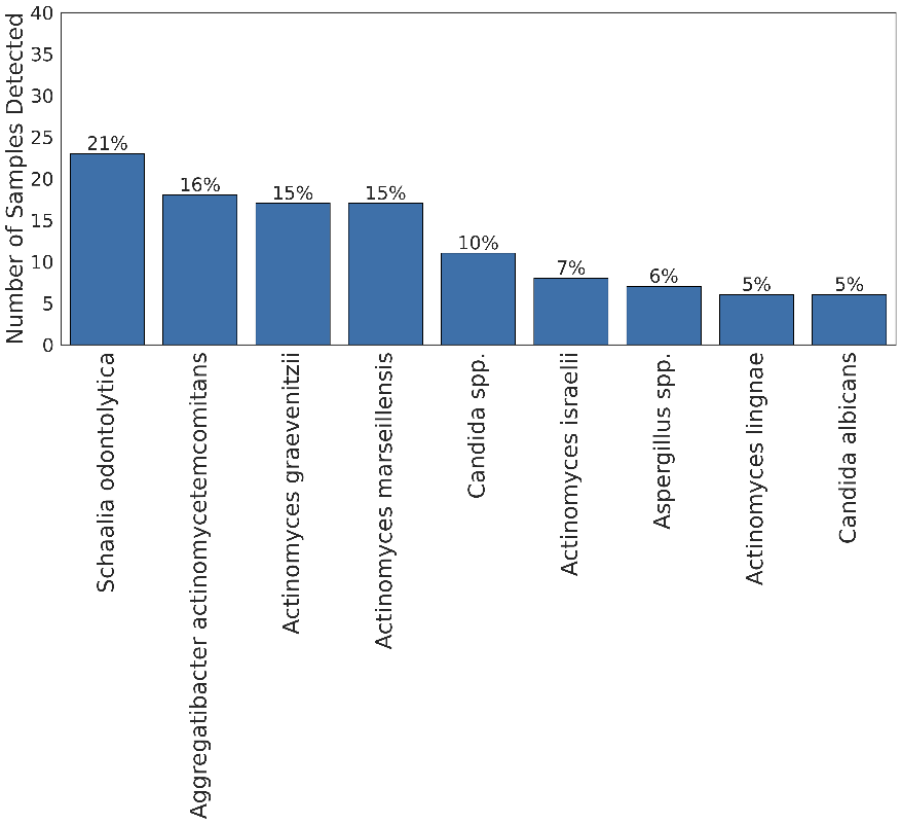


**Figure 3.** ***Results for the Hybrid Capture Screening of BAL.*** **A)** The distribution of the number of species detected per BAL sample. **B)** The organisms detected in at least 5% of the screened BAL samples.

## *In-silico* Validation of the Hybrid Capture Assay

An *in-silico* validation method was devised to test the ability of hybrid capture to properly classify organisms that were unable to be validated at the bench and whose gene targets share at least 95% sequence identity. Organisms present in each cluster were then paired and public sequencing datasets were obtained for each organism. These data sets were then mixed at relevant read counts based on the average number of on-target reads obtained from the validation data. The artificially generated datasets were analyzed using the custom bioinformatic pipeline to determine its ability to detect each organism in the paired datasets (Figure 4A).

Overall, 140 species were paired and correctly classified by the analysis pipeline (90% of tested species) (Figure 4B). 15 species were either not detected or incorrectly classified, all of which were removed from the target organism list. These *in-silico* studies provide confidence in the hybrid capture analysis pipeline's ability to properly distinguish and classify closely related species.

**Figure 4A**

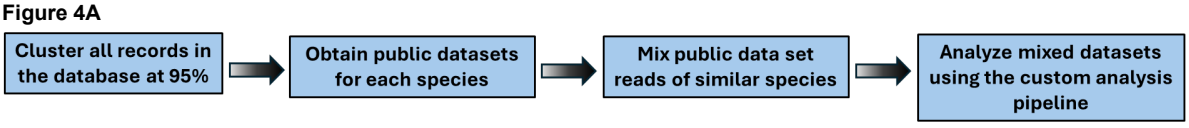

Figure 4B

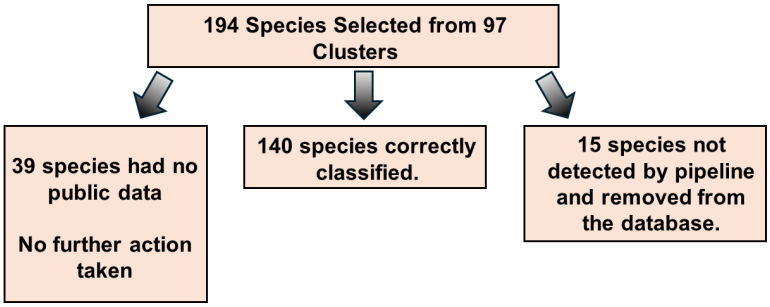


**Figure 4.** ***In-silico Validation of the Hybrid Capture Assay*** **A)** Overview of the *in-silico* validation **B)** Summary of the *in-silico* validation results.

## Conclusions

Next generation sequencing based diagnostics offer a powerful tool in the diagnosis of IFBIs., Targeted sequencing methods, such as hybrid capture, look to overcome the shortcomings associated with current whole genome methods, e.g, low sensitivity, low specificity, and complex reporting[5–7]. The assay described in the present study takes a hybrid capture approach to specifically target fungi and bacteria of clinical interest. Analytical accuracy and specificity validation data demonstrate the robustness of the cutoff criteria established for filtering out false-positive results with minimal effect on sensitivity. Furthermore, the presented hybrid capture assay shows similar sensitivity compared to reference laboratory qPCR methods, as demonstrated by the clinical accuracy dataset. Finally, a key benefit of hybrid capture is its targeted, yet still broad, nature. By targeting individual pathogens, hybrid capture can dedicate more sequencing reads to clinically relevant pathogens resulting in a low number of difficult to interpret polymicrobial results.

The data presented here demonstrates the key advantages of a hybrid capture diagnostic in diagnosing IFBIs. Future clinical studies will help bolster the analytical validation described in this report. Furthermore, development of additional hybrid capture probe panels will expand the current assay's capabilities to one day allow for a personalized approach to IFBI diagnosis.

## References


1. Brown, G. D., dW, D. D., IA G, N. R., tM, S. L. & Netea, M. G. Hidden Killers: Human Fungal Infections. www.ScienceTranslational Medicine.org.
2. Wickes, B. L. & Wiederhold, N. P. Molecular diagnostics in medical mycology. Nat Commun 9, 5135 (2018).
3. Blauwkamp, T. A. et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nat Microbiol 4, 663–674 (2019).
4. Samuel, L. Direct-from-Blood Detection of Pathogens: a Review of Technology and Challenges. Journal of Clinical Microbiology vol. 61 Preprint at https://doi.org/10.1128/jcm.00231-21 (2023).
5. Vissichelli, N. C. et al. Cell-free next-generation sequencing impacts diagnosis and antimicrobial therapy in immunocompromised hosts: A retrospective study. Transplant Infectious Disease 25, (2023).
6. Kaur, I. et al. Real-world clinical impact of plasma cell-free DNA metagenomic next-generation sequencing assay. Infect Control Hosp Epidemiol 1–8 (2025) doi:10.1017/ice.2024.242.
7. Edward, P. & Handel, A. S. Metagenomic Next-Generation Sequencing for Infectious Disease Diagnosis: A Review of the Literature with a Focus on Pediatrics. J Pediatric Infect Dis Soc 10, S71–S77 (2021).
8. Guitor, A. K. et al. Capturing the resistome: a targeted capture method to reveal antibiotic resistance determinants in metagenomes. Antimicrob Agents Chemother 64, (2019).
9. Cui, L., Morris, A. & Ghedin, E. The human mycobiome in health and disease. Genome Med (2013).
10. Könönen, E. & Wade, W. G. Actinomyces and related organisms in human infections. Clin Microbiol Rev 28, 419–442 (2015).


View Our Test Menu

# Get Fast Accurate Results

**Find out more at eurofins-viracor.com**
**or contact us at info@eurofins-viracor.com**
**or call 800-305-5198**

## About Viracor

With over 40 years of diagnostic expertise in infectious disease, immunology and allergy testing for immunocompromised and critical patients, Eurofins Viracor is passionate about delivering accurate, timely and actionable results, never losing sight of the connection between the testing it performs and the patients it serves.

Eurofins Viracor is a subsidiary of Eurofins Scientific (EUFI.PA), a global leader in bio-analytical testing, and one of the world leaders in genomic services. For more information, please visit eurofins.com and eurofins-viracor.com