Cross-Contamination in NGS Oligo Production: The Hidden Variable That Can Undermine Your Sequencing Data

Next-generation sequencing (NGS) workflows have become increasingly sophisticated. Researchers invest heavily in optimising sample preparation, sequencing platforms, bioinformatics pipelines and experimental design. Yet one critical variable often receives far less attention than it deserves: the quality and purity of the oligonucleotides used to build sequencing libraries.

Even low levels of cross-contamination between adapters and indexing primers can introduce sequencing artefacts that are difficult to identify, troubleshoot or correct once sequencing has been completed. As sequencing applications become increasingly sensitive and quantitative, controlling cross-contamination during oligo manufacturing has become a fundamental requirement for generating reliable data.

"As sequencing applications become increasingly sensitive and quantitative, controlling cross-contamination during oligo manufacturing has become a fundamental requirement"

Current Threshold Expectations in NGS Adapter and Index Oligo Manufacturing

A cross-contamination level of <0.1% for NGS adapters and indexing oligos is generally accepted across the industry because it represents a practical balance between sequencing accuracy, manufacturing capability and the error rates inherent in sequencing platforms. However, this is an industry convention rather than a hard scientific limit.

Index misassignment is already non-zero

Even if the oligos were theoretically 100% perfect, sequencing instruments introduce some degree of index misassignment, often referred to as index hopping or barcode swapping.

- Depending on the platform and library preparation method, this background can range from <0.01% to around 1%.
- Oligo contamination below 0.1% is typically low enough that it does not become the dominant source of incorrect sample assignment.

Manufacturing limitations

Oligonucleotides are synthesised on shared equipment.

- Despite rigorous cleaning procedures and dedicated workflows, completely eliminating every trace of a previous sequence is technically challenging.
- Modern quality-control methods can reliably demonstrate contamination below approximately 0.1%, making it an achievable specification for routine production.

Minimal impact for most multiplexed experiments

Consider a pool of 96 equally represented samples.

- If an index contains 0.1% contamination from another index, approximately one in 1,000 molecules could potentially receive the wrong barcode due solely to the contaminant.
- For many applications—including RNA-seq, whole-genome sequencing, ChIP-seq and amplicon sequencing—this is well below the level likely to affect biological conclusions.

Below typical QC detection limits

- Detecting contamination substantially below 0.1% often requires highly sensitive analytical methods.
- Routine QC methods such as LC-MS and capillary electrophoresis can confirm oligo identity and purity but do not directly quantify sequence cross-contamination at extremely low levels.
- The approaches required to verify contamination below 0.1% are more specialised, time consuming and costly.

Historical industry standard

The introduction of patterned flow cell platforms in 2017 and the subsequent industry concern generated where elevated rates of index misassignment highlighted the importance of both robust indexing strategies and stringent oligo manufacturing controls. Many oligo manufacturers and sequencing users have since adopted <0.1% as a specification because practical experience has shown it delivers reliable multiplexed sequencing performance for the vast majority of applications.

Why 0.1% Became the “Sweet Spot”

Threshold	Benefit	Cost
<1%	Easy to manufacture	Too many sample misassignment
<0.1%	Negligible impact for many NGS applications	Routinely achievable with robust manufacturing controls
<0.01%	Required for ultra-sensitive applications	Significantly higher manufacturing and QC requirements and costs

As a result, <0.1% has become a widely accepted industry benchmark: low enough that manufacturing-derived contamination is typically outweighed by other sequencing error sources, while remaining practical and economical to achieve. It is not a universal scientific threshold, but rather an engineering specification that has proven effective for standard NGS workflows.

"The accepted <0.1% contamination threshold is not a universal scientific limit, but an engineering specification that has proven effective for standard NGS workflows."

How Do NGS Oligo Providers Validate Their Cross-Contamination Rates?

Understanding the accepted contamination threshold naturally raises another question: how do manufacturers demonstrate that those levels are actually being achieved?

The key point is that routine QC methods generally do not directly measure sequence cross-contamination down to 0.1%. Performing comprehensive cross-contamination analysis on every production lot would add both cost and time to manufacturing.

Instead, manufacturers typically combine validated manufacturing processes with routine analytical testing to demonstrate that contamination remains below their target threshold.

Common approaches include:

- Validation of manufacturing processes
- Routine QC testing of mass, purity, concentration and identity on every batch
- Periodic cross-contamination analysis on representative lots
- Additional testing when introducing new equipment, cleaning procedures or production workflows

The methods and testing frequency vary between suppliers. If validation consistently demonstrates carryover below the accepted threshold, routine production can rely on those validated procedures.

This approach is similar to pharmaceutical manufacturing: not every tablet is tested for every conceivable impurity. Instead, the manufacturing process is validated to consistently produce material that meets predefined specifications, with periodic verification.

For this reason, manufacturer claims of contamination below 0.1% generally refer to process validation supported by highly sensitive contamination verification studies rather than routine LC-MS or HPLC testing alone.

Understanding Cross-Contamination in Oligo Production

Cross-contamination occurs when small amounts of one oligonucleotide sequence are unintentionally introduced into another during manufacturing, purification, handling or dispensing. While contamination levels may be very low, modern sequencing technologies can detect signals at correspondingly low frequencies.

Potential sources of contamination include:

- Residual phosphoamidites in synthesis systems
- Inadequate cleaning between synthesis runs
- Purification carryover
- Aerosol generation during liquid handling
- Manual handling errors
- Plate reformatting and dispensing processes

Because NGS workflows routinely amplify DNA through PCR and generate millions of reads, even trace contamination can become visible in the final dataset.

Why Traditional Oligo Quality Control Is Not Always Enough

Historically, oligo quality has been assessed using methods such as UV absorbance, mass spectrometry and HPLC analysis. These techniques remain essential for confirming oligo identity and purity, but they were not designed to detect low-level sequence contamination between closely related oligos.

This creates an important distinction between:

Chemical purity – confirming that an oligo was synthesised correctly.

Sequence purity – confirming that unintended oligos are not present at experimentally significant levels.

For many NGS applications, sequence purity may be the more important metric.

An oligo can meet traditional QC specifications while still containing contaminating sequences at frequencies capable of affecting sensitive sequencing experiments.

While process validation, routine QC and sensitive contamination monitoring demonstrate that a manufacturing process can consistently produce acceptable material, they cannot completely eliminate the possibility of isolated contamination events resulting from operator error, equipment malfunction, reagent mix-ups or unexpected process deviations.

The Experimental Consequences

1. Sample Misassignment

One of the most immediate consequences of oligo contamination is the incorrect assignment of sequencing reads to samples.

This can generate misleading results, particularly in highly multiplexed sequencing runs where large numbers of samples are pooled together.

Potential consequences include:

- Reduced confidence in sample identity
- Increased background noise
- Loss of low-frequency signals
- Additional data-cleaning requirements

In severe cases, experiments may need to be repeated entirely.

One of the most widely recognised sources of sample misassignment is index hopping, in which library fragments acquire different index sequences during cluster generation or sequencing. While modern sequencing platforms and library preparation methods have significantly reduced this effect, it remains an important consideration in multiplexed experiments.

Cross-contamination during oligo manufacturing and index hopping are distinct phenomena, but their consequences can appear remarkably similar. Both can produce incorrectly assigned reads and contribute to low-level background signals that complicate data interpretation. Researchers may observe unexpected variants or rare sequences without immediately knowing whether they originated during sequencing or earlier in the workflow.

"The smaller the signal being measured, the greater the potential impact."

2. False-Positive Variant Detection

Many sequencing applications involve detecting rare variants, mutations or low-abundance genetic signatures.

If contaminating oligos introduce reads originating from another sample, those reads may be interpreted as genuine biological signals rather than manufacturing artefacts. This can create the appearance of variants that do not actually exist in the sample being analysed.

For researchers studying:

- Oncology biomarkers
- Liquid biopsy samples
- Microbial diversity
- Somatic mutation detection
- Population genetics

even a small number of contaminating reads can influence conclusions.

3. Distorted Quantitative Measurements

Modern sequencing has increasingly become a quantitative technology rather than a purely qualitative one.

Researchers often rely on read counts to estimate:

- Variant frequencies
- Gene abundance
- Species abundance
- Molecular complexity
- Expression levels

Cross-contamination can skew these measurements by introducing molecules that artificially inflate or suppress observed frequencies.

The smaller the signal being measured, the greater the potential impact.

4. Increased Background Noise

Background noise is the enemy of sensitivity.

When contaminating sequences are present, they compete with genuine reads and make it more difficult to distinguish meaningful biological signals from technical artefacts. This can reduce:

- Detection sensitivity
- Statistical confidence
- Reproducibility
- Experimental power

Researchers may compensate by increasing sequencing depth, but this increases costs without necessarily solving the underlying problem.

This is one reason why high-quality index design is so important. The adoption of Unique Dual Indexes (UDIs) has become increasingly common because a read is assigned to a sample only when both index sequences match the expected combination. A single index-hopping event is therefore insufficient to produce a valid sample assignment, dramatically reducing the risk of misattributed reads.

However, even the most robust indexing strategy cannot compensate for contamination already present within the indexing reagents themselves. This highlights an important principle: effective sample assignment depends on both intelligent index design and stringent oligo manufacturing controls. High-quality oligo production and careful index design should be viewed as complementary safeguards rather than alternative solutions.

"Even the most robust indexing strategy cannot compensate for contamination already present within the indexing reagents themselves"

Why the Risk Is Growing

As sequencing technologies improve, analytical sensitivity continues to increase.

Applications that once operated comfortably above contamination thresholds are now routinely measuring signals that were previously undetectable. Researchers increasingly expect reliable detection of rare events, low-abundance molecules and subtle biological differences.

At the same time, sequencing facilities are processing larger sample batches than ever before. Higher levels of multiplexing place additional demands on adapter design, index quality and manufacturing consistency. Under these conditions, even low-level contamination can become experimentally significant.

Contamination Control Must Be Designed into Manufacturing

The most effective approach to contamination control is prevention rather than correction.

This requires manufacturers to evaluate contamination risk throughout the entire production workflow, including:

- Synthesis platform design
- Equipment cleaning procedures
- Purification workflows
- Liquid handling processes
- Operator handling protocols
- Automated dispensing systems
- Batch release testing

Importantly, contamination control should not be viewed as a final QC checkpoint. It should be treated as an intrinsic property of the entire manufacturing process.

"Contamination control should not be viewed as a final QC checkpoint. It should be treated as an intrinsic property of the entire manufacturing process"

The Role of NGS-Based Quality Control in Oligo Production

For sequencing applications, sequencing-based QC provides a particularly powerful solution.

By directly sequencing manufactured oligos, suppliers can quantify cross-contamination at levels relevant to end users. Unlike conventional QC methods, NGS-based testing can provide direct evidence of whether neighbouring sequences are present within a product batch and at what frequency.

This aligns the quality-control method with the intended application itself, making it especially valuable for high-sensitivity sequencing workflows.

The economic case for routine batch-specific NGS testing is also compelling. Although sequencing every batch increases both manufacturing time and cost, these expenses are often negligible compared with the scientific and financial consequences of releasing contaminated adapters or indexing primers to end users.

A single failed sequencing experiment can represent thousands to tens of thousands of pounds, euros or dollars in library preparation reagents, sequencing instrument time, bioinformatics analysis and personnel effort costs. Indirect costs—including project delays, repeat sample collection, lost productivity and missed deadlines—can be even greater. For irreplaceable samples such as patient biopsies, rare clinical specimens or longitudinal study materials, failed experiments may be impossible to repeat.

"The economic case for routine batch-specific NGS testing is also compelling"

Addressing the Challenge: How Biolegio Redesigned Its NGS Oligo Manufacturing Pipeline

Recognising the growing importance of sequence purity in modern sequencing workflows, Biolegio spent recent years re-evaluating and redesigning its NGS oligo manufacturing process with cross-contamination control as a primary objective.

Biolegio began with a comprehensive risk assessment of the entire production workflow, examining potential contamination sources across synthesis, purification, liquid handling, plate reformatting, dispensing and operator interactions. The analysis demonstrated that contamination risk was not confined to a single manufacturing step but existed throughout the process.

As a result, contamination control became a key design criterion across the production pipeline. Synthesis conditions, handling procedures, purification workflows and dispensing operations were reviewed with the specific goal of minimising opportunities for sequence carryover and inter-well contamination. New standard operating procedures were developed specifically for NGS adapter and indexing products.

The company also invested in additional automation and dedicated manufacturing infrastructure. Automated liquid handling was introduced where manual processing represented a contamination risk, improving consistency while reducing opportunities for handling-related errors. In parallel, Biolegio implemented synthesis platforms specifically configured for NGS oligo production with contamination control incorporated into the system design itself.

Perhaps most significantly, Biolegio developed a dedicated NGS-based quality-control workflow capable of directly measuring cross-contamination through sequencing. Unlike traditional analytical methods, this approach enables contaminating sequences to be detected and quantified at frequencies relevant to modern sequencing applications. The resulting data provide batch-level evidence of sequence purity and help identify contamination events that may occur during manufacturing.

Importantly, Biolegio regards this testing as part of a continuous improvement cycle rather than simply a product-release requirement. Sequencing-based contamination data are used not only to qualify products but also to inform ongoing process optimisation.

By redesigning its NGS oligo manufacturing pipeline around contamination control, Biolegio has also been able to exert greater control over the time and cost associated with enhanced quality assurance. This has allowed the company to maintain pricing for predefined adapter and indexing primer plates within expected market ranges while providing additional quality assurance without transferring significant additional costs to customers.

The ultimate goal is straightforward: to ensure researchers can be confident that any signal detected in their sequencing data originates from the sample itself rather than from the reagents used to prepare it. As sequencing applications continue to demand greater sensitivity and quantitative precision, this level of manufacturing control is becoming increasingly important.

Supporting this objective, Biolegio now applies a QC clearance threshold of 0.01% cross-contamination for all predefined adapter and index oligo products

Conclusion

Cross-contamination in NGS oligo production is rarely visible, but its effects can be far-reaching. From sample misassignment and false-positive variant calls to distorted quantitative measurements and reduced analytical sensitivity, contamination can compromise confidence in sequencing data long before bioinformatics analysis begins.

As sequencing technologies continue to evolve, researchers should evaluate oligo quality not only in terms of chemical purity, but also in terms of demonstrated sequence purity and contamination control. Ultimately, the reliability of sequencing data depends on the integrity of every component within the workflow—including the oligos used to build the library itself.

Biolegio's LegiSeq NGS Ready-to-Use Oligos are designed and manufactured to meet exactly these requirements — including NGS-verified cross-contamination rates at <0.01% for every batch, optimally designed 10-base UDIs, and UMI-incorporated adapter architectures. Manufactured under ISO 13485, every product leaves our facility with the quality documentation your workflow demands.