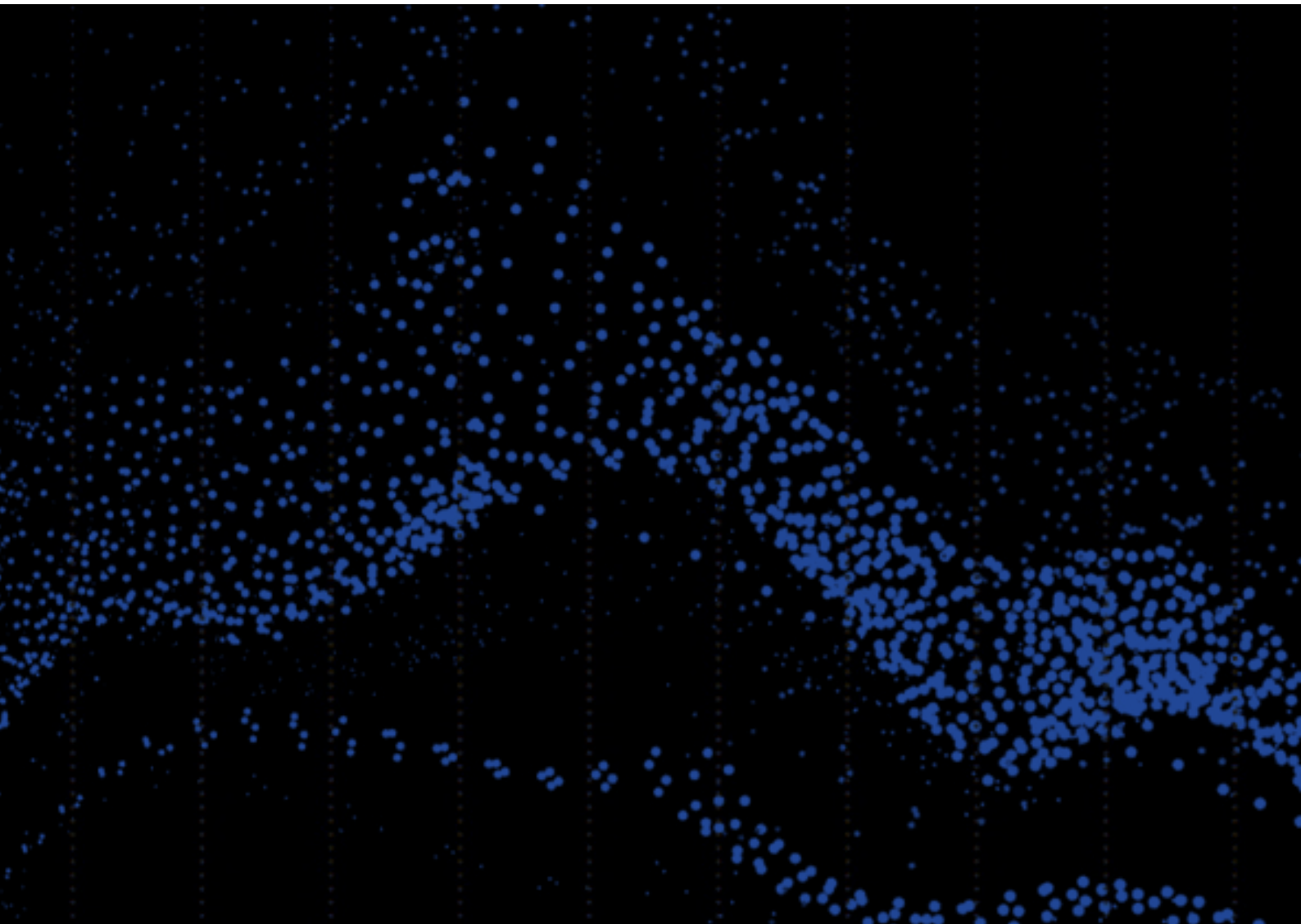
univ8 GENOMICS | EuroClonality

EuroClonality-NDC Assay

Analysis Guide

Release: February 2024

For **Research** Use Only. Not for use in diagnostic procedures.



EuroClonality-NDC Assay

This document is intended to guide the users through the analysis of the results generated by the EuroClonality-NDC assay and associated analysis tools[1,2]. Although every effort has been made to keep the information correct and up to date, we cannot guarantee that there are no unwilling errors or omissions. We encourage our customers to provide feedback which may lead to modifications and improvements in future versions of this document.

Warranties

No warranties are provided, explicitly or implied and Univ8 Genomics does not accept any liability for any consequences, direct or indirect arising from the use of the EuroClonality-NDC bioinformatics or analytical tools.

Trademarks

All trademarks are the property of their respective owners.

Copyright

© 2024 Univ8 Genomics Ltd. All rights reserved.

Univ8® is a Registered Trade Mark of Univ8 Genomics Limited.

Table of Contents

Introduction

Intended Use

The EuroClonality-NDC is a capture-based targeted next generation sequencing (NGS) assay with customised bioinformatics pipeline for research use only. It has been designed and analytically validated for the integrated and qualitative detection of clonal immunoglobulin (IG) and T cell receptor (TR) rearrangements, translocations, copy number variation (CNV) and somatic mutations from suspected cases of lymphoproliferative disorders or lymphoid populations. The EuroClonality-NDC pipeline has been validated for use with genomic DNA (gDNA) isolated from fresh/frozen and formalin fixed paraffin embedded (FFPE) tissue. However, the CNV tool has been validated only for high molecular-weight DNA from fresh specimens such as peripheral blood, bone marrow or separated cells.

Accessing the EuroClonality-NDC analysis application

The "EuroClonality-NDC analysis" application is now available on Illumina's BaseSpace Sequence Hub (BSSH) exclusively for EuroClonality-NDC customers. BSSH provides a secure portal for genomic data analysis and storage, with servers across multiple locations to support the requirements of GDPR, CLIA, HIPPA, multiple ISO standards and others. For more information, please refer to Illumina's BSSH help centre on the following link (https://help.basespace.illumina.com). For a detailed description of the security and compliance features of Illumina's BSSH, please follow these links (https://www.illumina.com/content/dam/illumina-marketing/documents/products/flyers/7139_AWS_Illumina_Solution%20Brief_SecurityCompliance_FINAL_20190405.pdf),(https://emea.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/basespace-security-and-privacy-security-brief-m-gl-01959/basespace-security-and-privacy-security-brief-m-gl-01959.pdf).

To use the EuroClonality-NDC analysis application you will first need to contact your Illumina local representative to open a Professional BSSH account for your team. This account will provide iCredits needed to run the EuroClonality-NDC analysis application. The computing cost of running the application is 3 iCredits per node/hour; this corresponds approximately to 1 iCredit per sample. A BSSH Professional account allows teams to manage access for their members and collaborators, in a secure and controlled fashion. Moreover, access can be customised to different levels for administrators, power users, curators or basic users, helping to implement local security policies as appropriate. The BSSH Professional accounts can be set up in different domains (e.g. EU, US, AUS, etc.), to ensure country-specific data protection compliance (e.g. GDPR), each of them with one "Working Group" (e.g. *A-Team EU*).

Once the BSSH Professional account has been set up, please follow the instructions below to access and run the EuroClonality-NDC application:

1. Log onto your regional domain for your team's specific BSSH Professional account with your personal account's username and password (e.g. for EU use https://euc1.sh.basespace.illumina.com).
2. On the top right corner, click on your username and select the Working group for your Professional account (e.g. *A-Team EU*).
3. From the top navigation menu, go to "APPS".
4. From the drop-down menu select *All Categories* or *Targeted Sequencing*.
5. On the "Search Apps" field, enter "Univ8" and click the search icon to see all the Univ8 Genomics apps.
6. Click to select the *EuroClonality-NDC Analysis App*. **Note:** you can bookmark all the Univ8 Apps for easier access in the future.
7. Launch the *EuroClonality-NDC Analysis App* by clicking on the "Launch Application" button.
8. Select the "Samples" option if you want to run samples from a single run (recommended and most common) or "Biosamples" option if you have run the same sample in multiple runs and wish to aggregate the data (not recommended).

9. Provide an *Analysis Name* (or use the generic default with datestamp).
10. Click on 'SELECT SAMPLE(S)' and choose which samples to run by ticking the corresponding boxes and then click "SELECT" to exit. **Note:** you can filter by "Project" to help with sample selection.
11. Click on "SELECT PROJECT" and select an existing project or enter new one.
12. After checking that everything is correct, click on "LAUNCH APPLICATION".
13. The *EuroClonality-NDC App* will be launched, and the samples will be run, wait until analysis has completed, approximately 20 minutes per sample (an email will be sent to you or the nominated admin account).
14. Download the data to your local server/computer by going to the corresponding "ANALYSIS", clicking on the "File" icon and selecting "DOWNLOAD">"ANALYSIS".
15. Make sure you install the "BaseSpace Sequence Hub Downloader" the first time (or when a new version is available) and then select "All file types including VCF, BAM, & FASTQ" and click "DOWNLOAD".
16. Select the local directory where you want to download the data into and continue.
17. Wait until the download has finished and check for integrity of the data in the corresponding directory.

Users must note that the FASTQ files will not be downloaded as part of the above download instructions and, if the users wish to keep these archived locally, they will need to download from the corresponding area (e.g. from "ANALYSIS" if they have been generated using Illumina's FASTQ generation App).

No data is ever transmitted to or accessible by Univ8 Genomics Ltd. or any third party, and the users are always in full control of their data, which they can choose to store, archive and/or delete from Illumina's BSSH at any time to comply with local security policies. Also note that storing data on Illumina's BSSH has a minimal overhead cost, but this can become significant if large data are stored for long periods of time (please contact Illumina support for detailed and up to date information).

Users must note that Univ8 Genomics Ltd does not warrant in any way, either explicitly or implied, the regulatory security compliance of the data transmitted or stored using the BSSH platform, and Univ8 Genomics Ltd does not accept any liability for any consequences, direct or indirect, arising from the use of the BSSH platform.

File Descriptions

EuroClonality-NDC analysis output files

The downloaded directory will contain one folder per sample with a series of outputs (file name/extensions in bold):

gathered.tsv & gathered.xlsx: this is the main output for analysis of the EuroClonality-NDC assay and shows the gathered results for IG/TR rearrangements (R), structural variants (SV) and sequence variants (VAR) in a single integrated file. Files are available in tab-separated-value (.tsv) and Microsoft Excel (.xlsx) formats for users to choose according to their downstream needs.

gathered.metrics.tsv: summary metrics per sample (e.g. percentage of duplicate reads or percentage of bases on/near target) as well as some EuroClonality-NDC-specific metrics.

bam & bai: alignment files and their indices required to visualise the data on any genome viewer such as IGV.

complete.vcf: VCF file (v4.3) containing the complete list of sequence variants (VAR) and corresponding annotation. All variants are included here (after removing known artefacts), irrespective of location, functionality, or population frequency.

curated.vcf: VCF file (v4.3) curated to remove artefacts, most non-functional variants as well as most of the common polymorphic variants and synonymous variants.

CNV.*.html: reports.zip: compress file containing copy number data files. After unzipping, the CNV.*.html files and corresponding libraries will appear containing CNV plots for IG, TR and cancer related genes (onco.html) for each sample. For each CNV plot there is a choice of FFPE or HMW, and users are encouraged to use the one corresponding to the type of sample analysed. Only HMW has been analytically validated for CNV analysis.

log: summary of the run and analytical output per sample.

Overview of the EuroClonality-NDC 'gathered' files

These files are the main output of the EuroClonality-NDC assay and contain information on all events apart from CNV. The following data items are included in the 'gathered' files:

name	R example	SV example	VAR example	description
sample	sample1	sample2	sample3	sample name as per FASTQ file
in FASTQs	6986785	12169894	3768296	total reads in sample FASTQs
duplicated	30.97%	34.38%	34.06%	percentage of reads duplicated
mapped	98.91%	98.88%	98.72%	percentage of reads mapped
on/near target	78.20%	76.21%	79.65%	percentage of bases on or near target
usable[a]	1599	2671	514	sum of fragments with rearrangement events
class1	R	SV	VAR	event class, R = rearrangement, SV = structural variant, VAR = sequence variant
class2	VJ:Vh-(Dh)-Jh	translocation	SNV	junction class (R) or variant type (SV/VAR)
locus	IGH			IG or TR locus (only for R)
gene	IGHV1-45	MYC	PIK3CA	gene involved in the event
gene partner	IGHJ6	IGHswc		partner gene involved in R/SV event
coord chr	14	8	3	chromosome containing event
coord pos	106506996	127736912	179234297	chromosomal coordinate
coord chr partner	14	14		partner chromosome
coord pos partner	105863195	105710469		partner chromosomal coordinate
% or ratio	19.0909	33.7657	28.9044	relative abundance, in fragments, in locus (R/SV) or variant allele frequency (for VAR)
numerator	21	755	124	number of fragments containing the event
denominator	110	2236	429	sum of fragments in locus (R) or total depth (SV/VAR)
in class	19.2661			relative abundance, in fragments, in class2 R (junction class)
of genes	95.4545			relative abundance, in fragments, in locus, of the clonotype genes involved in the R (event1)
event1	V1-45 J6	MYC / IGHswc	PIK3CA	clonotype genes (R) or genes involved in the SV (SV) or gene containing the variant (for VAR)
event2	-0/7/-4		c.3140A>G p.His1047Arg	clonotype segmentation (R) or HGVS (VAR); empty for SV
event3	CAR*GE#TS-DYGMDVW	t(8;14)	PASS;ECNDC.v2.PON=curated	clonotype aa junction sequence (R) or translocation (SV) or variant annotation category[b] (VAR)
event comments	[...][c]		[...][c]	functionality call and safety tags (R) or variant annotation (VAR); empty for SV
junction	CAR*GE#TS-DYGMDVW	8:127,736,912 / 14:105,710,469	H/R	clonotype aa junction sequence (R) or genomic breakpoint coordinates (SV) or aa change (VAR)
junction nt seq	tgtgcaagata GGGAGAG ....#ctacttctg	GCCTTATGAA TATATTCACGC TGACTCCCG	cAt/cGt	clonotype nt junction sequence (R) or clipped (30nt before/after) genomic breakpoint sequence (SV) or nt triplet change (VAR)
sequence context	[...][c]	[...][c]	A>G	clonotype nt sequence (R) or genomic breakpoint sequence (SV) or Ref/Alt change (VAR)

Shaded rows indicate the value may have different meaning, depending on the type of variant. [a] 'usable' is currently based on fragments with rearrangements, and is an indicator for quality and information content of the sample (see 'ANALYSIS: IG and TR rearrangement analysis' section below). [b] Variant annotation category contains the values available for 'event3' when analysing sequence variants (VAR): they provide a combination of overall quality (e.g. PASS, f0.02, etc) and whether the variant has been classified as 'curated' or 'complete' (see Analysis section). [c] These tend to be long strings of text, for examples please refer to the downloadable files at www.univ8genomics.com.

Analysis

IG and TR rearrangement (R) analysis

Each rearrangement (class1 = 'R') row in the 'gathered' files represents a specific clonotype for either a V-D-J, V-J, D-J, Kde, D-D or V-D event. The analysis provides all types of rearrangements, including complete, incomplete and even those that occur between different loci (i.e. TRA-TRD, which can also appear as SV).

As opposed to PCR-based NGS methods, where the V-D-J rearrangement sequences have the same start-end positions (i.e. primer locations), the EuroClonality-NDC data is generated upon hybridisation of a shotgun library preparation and, therefore, clonal fragments from different cells will have different start and end positions, making it more complex to consistently assign reads to clonotypes. This results in multiple clonotypes sometimes being assigned for an actual single clonal event. For example, a sample with a clonal TRGV4-TRGJ2 rearrangement, could have a number of entries for such rearrangement with different clonotypes (including 'X' if the junction is not fully sequenced), each of them with varying number of reads (i.e. 'numerator' in the 'gathered' files). Users should check the 'of genes' and 'event1' columns, which the report uses to provide information on combined frequencies, by assuming that all clonotypes with the same genes (e.g. TRGV4-TRGJ2) come from the same event, irrespective of the sequence.

Based on the EuroClonality-NDC validation performed by the EuroClonality-NGS working group, a minimum of 6 reads/fragments are required to assign clonality to a given rearrangement. This means that users can focus on analysing only events with 'numerator' ≥ 6 for the purposes of clonality. This value assumes similar unique depth as achieved during the validation work, which generally corresponds to 'unique mean target cover' values of between 500 and 1,000. Furthermore, this does not mean that events with <6 reads cannot be clonal. Conversely, events with ≥6 reads are not always high-quality clonal events. The users are encouraged to consider other relevant information provided in the gathered files and advice from this analysis section.

Thanks to the unique design and integrated analysis of the EuroClonality-NDC assay, we can estimate relative abundance of clonal rearrangements within each locus (e.g. IGK) and within each class (e.g. either Vk-Jk, Vk-Kde or intron-Kde). These values (i.e. '% or ratio' and 'of genes' for locus and 'in class' for class) provide valuable evidence on the likelihood that the events are clonal, and should be used alongside the total number of reads and clonotype checks.

For analysis of IG/TR rearrangement repertoires, where each rearrangement is expected to occur in only 1 or a few reads, no filtering should be performed, other than perhaps by 'class2', to omit non-functional/incomplete rearrangements., and by 'event3', to omit those with no junction region (i.e. 'X'). In cases where low-level clonal populations are expected (e.g. immuno-oncology studies or TIL analyses in solid tumours), particularly when analysing lymph nodes or any other lymphocyte-rich specimen, some rearrangements may be present in relatively high numbers, although as opposed to clonal lymphoproliferations, their relative abundance (i.e. '% or ratio') may be low.

Structural variant (SV) analysis

Analysis of SV is relatively simple as the number of events per sample with ≥ 6 fragments is generally low. Artefactual SV calls from background noise can sometimes occur with >6 reads. These are easily identified as the 'sequence context' column in the gathered file will likely contain highly repetitive sequences (e.g. homopolymer stretches >10bp long) or incomplete sequences containing several 'N' nucleotides.

SV can be confirmed visually using IGV or any other genome visualisation tool and we recommend performing this step in all instances. It can also help ascertain the exact makeup of the breakpoint and whether it is a balanced translocation or not.

The main limitation of SV analysis is the capture of at least one of the regions involved in the translocation. In particular, the EuroClonality-NDC can only detect SV in the regions described in the Quick Reference Guide. For example, MYC translocations can only be detected if the partner is included in the panel (e.g. IGH, IGK, IGL) or if the breakpoint in MYC falls within the region covered. Additionally, there are areas that are not fully covered or where the breakpoints may fall outside of the regions covered, in particular IGH switch regions or *KMT2A*, where up to 10-15% of breakpoints could be potentially missed.

Sequence variant (VAR) analysis

There are two main ways of analysing sequence variants with the EuroClonality-NDC assay, the recommended option is to analyse all variants in an integrated fashion using the 'gathered' files, however you can also analyse sequence variants only using the provided VCF files. There are two VCF v4.3 format files, one labelled 'complete' containing all variants (after removal of potential artefacts), and the other labelled 'curated' containing the more functionally relevant variants (see above under 'outputfiles' section). Both are annotated using VEP on hg38 (ensembl.org) and include the MANE Select transcript (or MANE Plus Clinical) when available. Note that some variants commonly known by their nomenclature using a previous transcript, could have changed when using MANE Select, e.g. the common hotspot variant in *MYD88* p.L265P is now reported as p.L252P with the current MANE Select transcript.

All variants included in the 'complete' VCF file are included in the 'gathered' files, and there is a flag under 'event3' that allows users to filter by only 'curated' variants. The variants with the 'curated' flag represent the most likely relevant variants for most analyses, including those that are functional, non-synonymous and present at low allele-frequency in the global population. Conversely, the variants with the 'complete' flag represent non-functional (or at least not clearly functional, such as intronic, regulatory, etc.), and/or synonymous and/or highly polymorphic variants. While the 'curated' flag is recommended for most applications the 'complete' flag can be used for specific applications or troubleshooting requiring 'identification' such as investigating contaminations, chimerism, or MRD and cfDNA analysis. The 'curated' variants include splice donor/acceptor sites within 2bp of the exon. If you need to analyse wider splice regions then use the 'complete' list. Note that the 'complete' VCF files contains all variants, while the 'complete' flag in the 'gathered' files contain only the non-'curated' variants.

Variants are reported here down to > 1.5% variant allele frequency (VAF), although the EuroClonality-NDC assay has been validated to report variants at > 4% VAF. Variants present at ≤4% VAF are likely the result of background noise and should not be analysed routinely. Nonetheless, the file provides these low-level variants for users to consider when expecting low-level variants or when analysing follow-up samples or known hotspot variants. Special caution must be taken with indel variants at low VAF, particularly those in highly repetitive regions, e.g. *KMT2D* or *RUNX1*, as these can represent artefacts, sometimes at VAF>4%. Variants in regions with low total coverage (e.g. <200x) should also be interpreted with caution.

Copy number variant (CNV) analysis

For each sample, there are six CNV plots in HTML format. These are designed for users to be able to perform CNV in high molecular weight (HMW) and formalin-fixed paraffin-embedded (FFPE) sample types. Performance metrics are only available for CNV in HMW samples (see below). The plots highlight each location in different colours, according to the likelihood of CNV being present: green are regions that may be deleted (logR: 0.7-0.8), blue regions are more likely deleted (logR < 0.7), orange represent potential gains (logR: 1.3-2) and red are areas likely to be amplified (logR >2). There are three CNV plots per sample type: IG, TR and Onco.

IG and TR: these plots display CNV in the IG and TR genes, respectively. Since the process of V-(D)-J recombination results in the deletion of the intervening genes, these CNV plots are useful to assess clonality and confirm the specific genes involved in the rearrangement (for examples of these plots visit www.univ8genomics.com).

Due to limitations of the capture-based NGS technology and the lack of some polymorphic IG/TR genes in the reference human genome, not all possible genes are displayed in these plots. Some have been taken out due to lower performance and some of these (but not all) may be associated with polymorphic variants. Some polymorphic variants, however, do remain in the CNV plots, which sometimes can cause confusion when visualising the plots - i.e. if the specimen contains the polymorphic loss. In those instances where the rearrangement involves genes not included in the display, the CNV plots should show the respective copy number loss of the genes up until the gene closest to the one reported in the 'gathered' files.

It is normal to see artefactual copy number gains in TR genes in cases with high number of IG rearrangements (either clonal or polyclonal) and vice versa, due to the way the baseline copy number values are generated for this analysis. It is of course possible that these gains or even amplifications can be seen in cases with aneuploidy in the respective chromosomes or chromosomal regions. Generally, however, gains of these regions are of no significant scientific or clinical relevance.

Onco: these plots display CNV in oncogenes and tumour suppressor genes, as well as some additional, clinically-relevant genomic locations included in the panel (e.g. 13qMDR, 11q, 17p, etc.).

A validation study carried out in HMW samples for four genomic regions (11q, 13q, chr12 and 17p) resulted in overall >95% sensitivity and specificity. This performance is limited for samples with CNV present in >20% of the DNA and therefore the recommended minimum tumour content is 40%.

The main limitations of the CNV analysis for cancer-related genes are: i) they have only been validated in HMW-DNA; ii) they can be highly affected by DNA quality and purity (ddH_2O elution is highly recommended); iii) CNV analysis is significantly affected by hyperdiploidy and CNV analysis should not be performed in samples with >20 regions altered; iv) under- or over-clustering can have a negative effect in CNV analysis (e.g. if using significantly more or less samples than recommended).

Analysis tips

VCF compatibility: the VCF files contain only sequence variant information in v4.3 format.

Integrated analysis: the 'gathered' files contain all the results apart from CNV. Users can filter and sort by any given header, to simplify analysis. It can be helpful to see all the types of variants together, in order to have a holistic view of the entire sample. Due to the large number of individual events, for suspected clonal specimens, this can be best achieved as follows:

- Filter your results by 'numerator' ≥ 6 to only show events supported by at least 6 fragments/reads (for higher coverage samples or if high clonality expected, this can be raised to >10 or even higher)
- Filter 'event3' column to omit any variant that contains the keyword 'complete' (this will show only 'curated' VAR, without filtering out R/SV events). Note: if using Microsoft Excel, use the filter tool and filter by 'does not contain' and enter the word "complete".

These settings will generally provide a list of approximately 20-40 variants per sample, including R, SV and VAR. Further refinement can be achieved by setting more stringent filters: e.g. 'selecting % or ratio' ≥4 will only show VAR with VAF≥4% and R and SV present in ≥4% in class, which reflects the limit of detection of the assay,) at the expense of potentially missing low-level or sub-clonal events.

False positive clonal R: due to the way capture-based data is analysed and the limited diversity of some genes, there are two main types of rearrangements that can lead to potential false positives being reported, these are IGHD7-27 and IGKde-IGKIntron rearrangements. The former is a result of the close proximity of IGHD7-27 to IGHJ1 and can result in sometimes reporting IGHD7-27 with IGHJ2 as a clonal rearrangement. In the case of Kde-Intron rearrangements, this is due to the limited diversity existing in this type of rearrangements. In both cases, if the junction amino acid sequence ('event3' column) has an 'X', or if the number of total reads and relative abundance is low, they can be safely assumed to be artefactual. No clonality calls should be made based on either of these rearrangements in isolation.

Other loci that can result in artefacts and require some analysis are:

Rearrangements involving TRBD1 and TRBD2 genes: due to their proximity to TRBJ genes and the low combinatorial possibilities, these events can frequently contain more than 6 reads. If there are several of them and the junctional regions ('event3') contain 'X' then they can be safely assumed to be artefactual.

TRG rearrangements: like the above, due to the low combinatorial possibilities and the preferential usage of certain TRGV and TRGJ genes, some artefactual events can be present with a significant number of reads. As above, if there are several of them and the junctional regions ('event3') contain 'X' then they can be safely assumed to be artefactual. Based on the large number of cases used during validation of the EuroClonality-NDC assay, most clonal T cell disorders will contain at least 1 or more TRA V-D-J and TRB V-D-J rearrangement in the 'gathered' file. No clonality calls should be made based on TRG rearrangements alone.

Absence of junction regions: for all IG and TR events, the lack of a junction region (i.e. 'X' in the 'event3' column) tends to be associated with artefactual rearrangements, even when the event has more than 6 supporting reads. In most cases these events can be safely assumed to be artefacts, unless there are other events in the 'gathered' file with the same rearrangement in more than 6 reads with a clear junctional region, in which case they are likely to represent the same clonal event.

Copy Number Analysis: CNV analysis is the most challenging part of the EuroClonality-NDC assay due to the complexity of including IG/TR regions, which naturally recombine in both normal and malignant samples with lymphoid content. These result in deletions of the intervening genes and can have significant variation on CNV profiles depending on type of rearrangement, lymphocyte content, etc. Even though CNV has been validated for HMW, achieving the expected performance may require substantial optimisation from lab to lab. Using purified cell populations (e.g. CD19+, CD138+, etc.) may also significantly help with CNV.

Example data files

Please visit the Resources section of www.univ8genomics.com to download example files from a number of cell lines analysed with the EuroClonality-NDC assay

Performance specifications

The validation and performance characterisation of the EuroClonality-NDC was carried out on 280 samples comprising a range of different B and T cell malignancies using a different version of library preparation and hybridisation methodologies. The EuroClonality-NDC assay and analysis workflows have been updated to improve performance and target coverage.

Overall performance of the EuroClonality-NDC assay was assessed by identifying underperforming regions using DNA extracted from a panel of 91 normal tissue samples. Regions were classed as underperforming if coverage was >2 standard deviations (SD) below the mean in ≥50% samples. Only 10 underperforming coding regions have been identified that could affect mutation detection in the genes and exons shown in the table below.

Gene Name	Chromosome	Exon(s)
ARID1A	chr1	6
CARD11	chr7	12
JAK2	chr9	15
TRAF2	chr9	1
PAX5	chr9	10
STAT5B	chr17	18
TCF3	chr19	7
BTK	chrX	18
PHF6	chrX	6
PHF6	chrX	3

Expected Metrics

The table below shows the approximate expected metrics of the EuroClonality-NDC assay based on 22 samples per hybridisation reaction (plus controls) and run on a NextSeq500/550 (Mid-Output 150 cycles cartridge) at 2 x 75bp. These can vary depending on DNA quality and quantity, technical factors in each laboratory, differences in equipment, etc.

Parameter	Expected HMW-DNA	Expected FFPE-DNA
Library yield (ng)	4,000 ng (range: 1,000-8,000)	
Library average fragment size (bp)	330 bp (range: 275 – 450 bp)	290 bp (range: 245 – 330 bp)
Average on/near target bases (%)	80% (range: 70 - 85%)	
Duplicate reads (%)	40% (range: 30 – 50%)	
Mean unique target coverage (reads)	1,000x (range: 500 – 1,500x)	

This improved version of the EuroClonality-NDC assay and workflow underwent a verification bridging study with gDNA extracted from 22 FFPE and 44 HMW samples which were included in the original validation study, and which comprised of a broad range of B and T cell malignancies. The following table represents the observed analytical performance:

	Sensitivity	Specificity	Limit of Detection
IG/TR Rearrangements	>97%	>97%	5% VAF
Translocation	>95%	>99%	5% VAF
SNV	>99%	>99%	4% VAF
CNV*	>95%	>97%	20%**

*SNV: single nucleotide variant; CNV: copy number variation. *CNV have been validated in HMW-DNA for trisomy 12 and deletions of 11q, 13q and 17p. No performance data available for FFPE or other genomic regions.**Deletions can be detected if present in ≥ 40% of the cells. The threshold baseline fold-change for deletions detected by EuroClonality-NDC is 20% (i.e. 0.8) for losses and 30% (i.e. 1.3) for gains, e.g. trisomy.*

Abbreviations

CNV	Copy number variation
FFPE	Formalin fixed paraffin embedded
HMW	High-molecular weight
IG	Immunoglobulin
SNV	Single nucleotide variant
TR	T cell receptor
SV	Structural variant

References

1. Stewart JP, Gazdova J, Darzentas N, Wren D, Proszek P, Fazio G, Songia S, Alcoceba M, Sarasquete ME, Villarese P, van der Klift MY, Heezen KC, McCafferty N, Pal K, Catherwood M, Kim CS, Srivastava S, Kroeze LI, Hodges E, Stamatopoulos K, Klapper W, Genuardi E, Ferrero S, van den Brand M, Cazzaniga G, Davi F, Sutton LA, Garcia-Sanz R, Groenen PJTA, Macintyre EA, Brüggemann M, Pott C, Langerak AW, Gonzalez D; EuroClonality-NGS Working Group. Validation of the EuroClonality-NGS DNA capture panel as an integrated genomic tool for lymphoproliferative disorders. Blood Adv. 2021 Aug 24;5(16):3188-3198. doi: 10.1182/bloodadvances.2020004056. PMID: 34424321.

2. Bystry V, Reigl T, Krejci A, Demko M, Hanakova B, Grioni A, Knecht H, Schlitt M, Dreger P, Sellner L, Herrmann D, Pingeon M, Boudjoghra M, Rijntjes J, Pott C, Langerak AW, Groenen PJTA, Davi F, Brüggemann M, Darzentas N; EuroClonality-NGS. ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data. Bioinformatics. 2017 Feb 1;33(3):435-437. doi: 10.1093/bioinformatics/btw634. PMID: 28172348.

Revision History

Release	Summary of change	Date
2100012 r01	First released	September 2021
2100012 r02	Updated references, updated FAQs and minor corrections	February 2024

Technical Assistance

Univ8 Genomics Ltd
63 University Road
Belfast
Northern Ireland

Email: contact@univ8genomics.com
Web: univ8genomics.com