CORVELVA
insieme dal 1993

Vaccinegate:
Metagenomic analysis report on Priorix Tetra

Introduction

As known, vaccines are biological drugs used to prevent certain infectious diseases. They are made up of several components: antigens (viruses, inactivated or attenuated bacteria, inactivated toxins, proteins or complex molecules derived from viruses and bacteria, able to stimulate the immune response), adjuvants (substances that increase the capacity of the vaccine antigens to induce the antibody immune response), excipients (substances needed to formulate the vaccine, or to preserve it from bacterial contamination) and contaminations (traces of substances derived from raw materials, eg cell lines for bacteria and virus growth, or from the production process, eg formaldehyde, antibiotics). During the registration phase of a biological drug, the vaccine is subject to the controls provided by the EMA guidelines and agreed with the regulatory institution according to the specific type of vaccine. These checks are then carried out on a representative number of samples on each Batch before marketing.

The responsible for the conformity of the product sold is therefore the manufacturer and the regulatory agencies in charge of control.

Since the safety of a vaccine depends on its compliance with the quality criteria, especially those regarding the absence of toxic or potentially toxic contamination (ie for which no effects in humans are known) it is of great importance that such compliance is strictly respected. Several studies in the literature have put the issue of the presence of various types of contaminations, both chemical and microbiological, thus opening the question if the vaccines actually comply with the directives imposed by the regulatory bodies, if in turn the regulatory agencies apply the control for the respect of these directives and if the regulatory agencies have defined with effective guidelines the criteria for the control and containment of such contaminations. To answer these questions, Corvelva commissioned the analysis of biological contaminations - which should never be present in vaccines- at a highly qualified center of services specialized in the genomic sequencing of DNA and RNA.

The study commissioned by Corvelva was based on two types of analysis:

1. **Testing of the presence of nucleic acids** (DNA / RNA) of human and animal origin and of microorganisms (viruses, bacteria) using the Next Generation Sequencing method, which allowed to quantify in a highly specific and accurate sequence the genetic material contained in the vaccines examined

2. **Verification of the correspondence of the genome sequences** of live attenuated or inactivated bacteria and viruses present in the vaccines (presence of genetic variants)

Description of the method used for the analysis

Next Generation Sequencing, also known as deep sequencing, generates a single sequence from each DNA fragment, or cDNA, present in a sample. The downstream bioinformatics analysis then allows the differentiation between the origin of the sequence fragments, for example human, bacterial species or a particular virus. This means that mixed biological samples can be easily solved with this technology, which has now entered the routine of genomic research and diagnostics. Moreover, from NGS data it is possible to reconstruct the entire sequence of viral DNA and RNA genomes and bacterial genomes present in the sample and compare it with the reference genomes present in public databases.The examined samples are shown below along with the results obtained, grouped by classes of similar vaccines:

** ssRNA: single strand RNA, single-stranded RNA; dsDNA: double strand DNA, double stranded DNA.*
The underlined terms are made up or contain genetic material (DNA and / or RNA)

CORVELVA
insieme dal 1993

Analyzed Batches

Batch #1 - A71CB205A and Batch #2 - A71CB256A

Product name:	**Priorix Tetra**
Type of product:	Tetravalent vaccine measles, mumps, rubella, varicella
Manufacturer:	GlaxoSmithKline, Belgium
Composition : [1]	live attenuated viruses: Measles (ssRNA) Swartz strain, grown in embryo chicken cell cultures; Mumps (ssRNA) strain RIT 4385, derived from the Jeryl Linn strain, grown in embryo chicken cell cultures; Rubella (ssRNA) Wistar RA 27/3 strain, grown in human diploid cells (MRC-5); Varicella (dsDNA) OKA strain, grown in human diploid cells (MRC-5)

Required analysis

Test the presence of nucleic acids (DNA / RNA) of human and animal origin and of microorganisms (viruses, bacteria), using a metagenomic / metatranscriptomic approach on the Illumina platform of Next Generation Sequencing.

From a comparison among the three vaccines it is possible to highlight the following critical issues:

Priorix Tetra is the vaccine with the highest amount of contaminating foreign DNA (Total DNA extracted from 3.7 µg to 1.7 µg, of which 88% is human, then coming from the MRC-5 cells, and the remaining 12% comes from adventitious microorganisms, such as viruses, bacteria, worms). Human genomic DNA has high molecular weight above 60,000 bp and the total sequential coverage of the entire human reference genome (HG-19) **shows that the entire genome of fetal cells used for the culture of vaccinia viruses is present and not just portions of it**.

From the EMA's answer to our question [2] about the limits imposed on residues of foreign genetic material in vaccines, it appears that in fact there are no limits for each vaccine but only for some, reported in the monographs of the product; **the maximum allowed limit ranges from 10 pg to 10 ng**, based on the theoretical calculation of the possibility for the foreign genomic DNA to cause oncogenic mutations.
It is noteworthy that regulatory authorities do not require these contaminations to be tested in the final product, but only in the initial preparation phase, and for the attenuated virus vaccines the purification of these contaminations is a critical step. [3] EMA has not provided specific studies on the dangers of fetal residual DNA, which allow assessing the risk to human health of these contaminations, so this limit remains arbitrary today.

It follows that for these two for these two batches of Priorix Tetra it is about 325 times higher than the maximum limit of 10 ng and 325,000 times higher than the minimum limit of 10 pg.

On the question of contaminating human DNA, the World Health Institute in an official 2011 document entitled 'Recommendations for the evaluation of animal cell cultures as substrate for the manufacture of biological medicine products and for the characterization of cell banks' argues that what is necessary to take into consideration with respect to rcDNA (residual cellular DNA) in vaccines is:

- **A.** reduction in the amount of contaminating DNA during the manufacturing process;
- **B.** reduction in the size of the contaminating DNA during the manufacturing process;
- **C.** chemical inactivation of the biological activity of DNA occurred during the manufacturing process.

Taking into account the three requests described above, the product is considered by their regulatory organs (NRA) and control laboratories (NLC) to be at an acceptable level of risk regarding the presence of DNA from the cell substrate, based on (a) and / or (b) and / or (c), when the data show that adequate levels of safety have been achieved.

In particular, in the 2 batches of Priorix Tetra vaccine tested to date, point A. does not occur because the quantity is about 140 times higher than that recommended by the FDA (in Briefing Document September 19, 2012: Vaccines and Related Biological Products Advisory Committee Meeting) and the EMA, ie ≤ 10ng per dose; point B. does not occur because the DNA is high molecular weight (most> 10,000 bp, as can easily be verified using a simple agarose gel to control the quality of the DNA extracted from the vaccine), ie 50 times greater than the size recommended by

[1] https://farmaci.agenziafarmaco.gov.it/aifa/servlet/PdfDownloadServlet?pdfFileName=footer_000200_038200_RCP.pdf&retry=0&sys=m0b1l3
[2] Quesito EMA: https://www.ivancatalano.eu/wp-content/uploads/2018/05/Letter.pdf
[3] http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003322.pdf
http://www.who.int/biologicals/Cell_Substrates_clean_version_18_April.pdf
http://www.who.int/biologicals/Molecular%20Methods%20Final%20Mtg%20Report%20April2005.pdf?ua=1
https://pdfs.semanticscholar.org/presentation/0be5/9fbe9c69baa35c0f24086936bc541809ebc6.pdf

the FDA (200bp or less). Finally, in the same vaccine, point C. does not occur because, containing attenuated viruses, a possible chemical DNA inactivation would also inactivate viruses.

Comparison between two batches of Priorix Tetra

Batch #1 - A71CB205A	Batch #2 - A71CB256A
DNA analysis **Total DNA extracted: 1.7 µg in total per dose of 0.5mL** DNA sequencing analysis performed using a metagenomic approach, out of a total of **3.830.074** sequences produced **Presence of genomic DNA of:** **Varicella 14%** **Chicken 4%** **Human (MRC-5) 74%** **proteobacteria 1%** **RNA virus 0.01%** **not assigned 5%**	**DNA analysis** **Total DNA extracted: 3.7 µg in total per dose of 0.5mL** DNA sequencing analysis performed using a metagenomic approach, out of a total of **5.836.297** sequences produced. **Presente DNA genomico di:** **Varicella 14%** **Human (MRC-5) 88%** **(about 3,3 µg that is equivalent to approximately 300.000 fetal cells)** **RNA virus 0.0003%** **not assigned 0.5%**
RNA analysis Total RNA extracted: not quantifiable through standard fluorimetric methods RNA-seq analysis performed using a metatranscriptomic approach, out of a total of 10.445.038 sequences produced. **Measles 0.004%** **Mumps 0.008%** **Rubella 0.00007%** **Varicella 5%** **other viruses approaximately 0.002%** **helminths 0.6%** **Chicken 0.2%** **Human 87%** **not assigned 5%**	**RNA analysis** Total RNA extracted: 200ng per dose da 0.5mL RNA-seq analysis performed using a metatranscriptomic approach, out of a total of 6.171.266 sequences produced. **Measles 0.004%** **Mumps 0.008%** **Rubella nd%*** **Varicella 7%** **other viruses approaximately 0.001%** **nematoda 1.50%** **Proteobacteria 5.5%** **Human 68%** **not assigned 6%** *** sequencing with 260.343.42 reads: 114 reads equal to 0.00004%**

CORVELVA
insieme dal 1993

Methods and results

DNA and RNA extraction Priorix Tetra batch A71CB205A

Batch **A71CB205A was processed in June 2018.**
Genomic DNA extraction was carried out with the Maxwell® 16 Blood DNA Purification Kit sold by Promega and with the automatic extractor Maxwell® 16 IVD (Promega), following the manufacturer's protocol.
RNA extraction was performed using the PureLinkTM Viral RNA / DNA Kit Mini Kit (Invitrogen) following the manufacturer's protocol.
The starting amount used for the extractions are as follows (starting from a single vial of product):

- Extraction DNA: 125 µl of the 500 µl of suspension for injection
- Extraction RNA: 125 µl of the 500 µl of suspension for injection

Extracted DNA quantification and quality control was performed with the Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA) and the NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA) respectively.
Following is the of the DNA quantifications (NanoDrop ND = 1000; QB = 2.0 Qubit; HS = dsDNA HS Assay Kit)

Sample ID	ND A260/280	ND A260/230	QB_HS_ng/µL	volume_µl	Tot_amount_ng
DNA lotA71CB205A	1.54	1.31	9.41	45	423.45
RNA lotA71CB205A*	n.d.	n.d.	Out of range	50	0

*Quantità di RNA sotto il limite di quantificazione del fluorimetro QBit.

DNA concentration measurement by QuBit fluorometer showed that **batch A71CB205A** contains a quantity of **gDNA of 1.7 µg total per dose of 0.5mL**, calculated as follows:
9.41ng/µl (concentration determined by QuBit) x 45 (final resuspension volume of DNA after extraction, expressed in microliters) x 4 (starting volume submitted to the extraction procedure that is 1/4 of the volume of the dose contained in the whole vial equal to 0.5mL).

DNA e RNA extraction Priorix Tetra batch A71CB256A

Batch **A71CB206A was processed in december 2018.** Some improvements in the procedure have been made, such as:

1. Use of a larger volume of starting injection solution in order to increase the amount of extracted RNA (in the extraction from the previous batch the amount of RNA obtained was below the detection threshold with fluorimeter);
2. We performed pulsed-field electrophoresis, in order to have a greater detail on the size of the whole genomic DNA present in the sample;
3. We used a most sensitive measurement mode for RNA (Agilent RNA 6000 Pico Kit on Bioanalyzer Agilent).

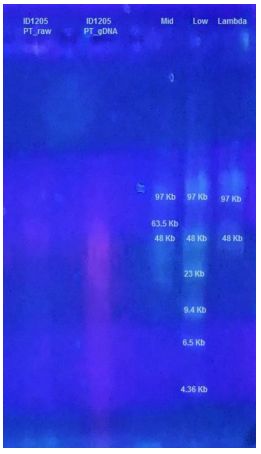
DNA extraction was performed by Maxwell® 16 Blood DNA Purification Kit resold by the Promega company and by automatic extraction (Maxwell® 16 IVD– (Promega) following the manufacturer's protocol.
RNA extraction was taken by PureLinkTM Viral RNA/DNA Mini Kit (Invitrogen) following the manufacturer's protocol.
The starting amount used for the extracions are as follows, starting from **two vials** of product of the same batch:

- DNA extraction: 300µl of the 500µl of suspension for injection (vial 1)
- RNA extraction: the entire volume of powder from a vial was resuspended in 200 µl instead of 500 µl of physiologic saline solution supplied in the package, and the entire volume was used for RNA extraction (vial 2).

Extracted DNA quantification and quality control was performed with the Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA) and the NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA) respectively.
Following is the of the DNA quantifications (NanoDrop ND = 1000; QB = 2.0 Qubit; HS = dsDNA HS Assay Kit)

Sample ID	ND A260/280	ND A260/230	QB_HS_ng/µL	volume_µl	Tot_amount_ng
DNA lotA71CB256A	1.95	2.12	40.8	55	2224

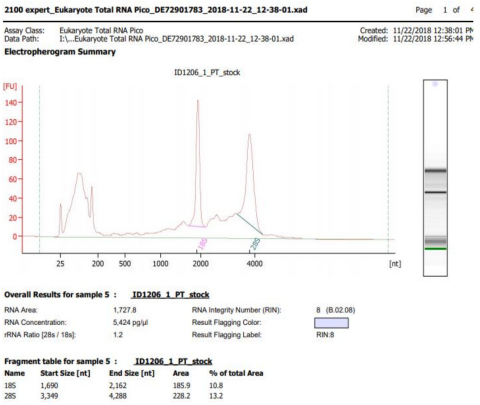
DNA concentration measurement by QuBit fluorimeter showed that **batch A71CB256A,** contains a quantity of **gDNA of 3.7 total per dose of 0.5mL**, calculated as follows:
40.8 ng/µl (concentration determined by QuBit) x 55 (final resuspension volume of DNA after extraction expressed in microliters) x 5/3 (starting volume submitted to textraction procedure that is 300 µl of the 500 µl of suspension).
Extracted DNA Pulsed-field electophoresis (PFGE, 5-80Kb, run time14h in TBE 0.5x, 80V) from batch A71CB256A, made visible thanks to SybrGreen fluorescent interlayer, showed the presence of a wide Genomic DNA 'strip' that reaches up to very high molecular weights but with a significant DNA amount in the 20-60Kbp range. In particular in the photo below, the ID1205_PT _raw sample is the vaccine contained material lysate before DNA purification, while PT_gDNA is the genomic DNA after extraction; Mid, Low and Lambda are 3 commercial molecular weight markers.



Extracted **RNA** quantification and quality control were performed by Agilent 2100 Bioanalyzer using Agilent RNA 6000 Pico Kit (Agilent Technologies, Santa Clara, CA). Below are the concentration values, the RNA (Integrity Number) measured at Bionalyzer:

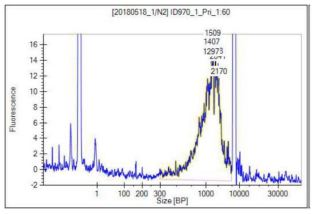
Sample ID	Bioanalyzer_pico_totale_ng/µL	RIN	volume_ul	Tot_amount_ng
RNA lotA71CB256A	5	8	37	200.688

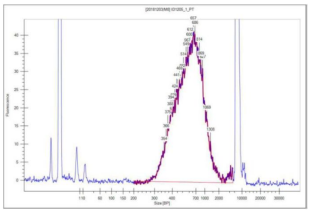
RNA amount contained in vaccine vial **batch A71CB256A was found to be about 200ng.**
The RIN equal to 8 indicates an of excellent quality RNA and an intact eukaryotic RNA, being present both the 18S and 28S peaks typical of eukaryotic RNA:



Preparing DNA-seq Library with Illumina technology

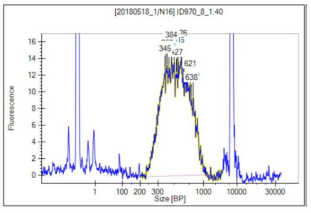
Kit Ovation® Ultralow System V4 1–96 (Nugen, San Carlos, CA) has been used to prepare the libraries according to the manufacturer's instructions, starting from 10ng genomic DNA. Final libraries have been quantified by fluorimeter Qubit 2.0 (Invitrogen, Carlsbad, CA) and quality-tested by the system Caliper GX (PerkinElmer, Waltham, MA) for batch A71CB205A and Agilent 2100 Bioanalyzer, DNA High Sensitivity Analysis kit (Agilent technologies, Santa Clara, CA) for batch A71CB256A. Set out below the tracing for the two obtained libraries:



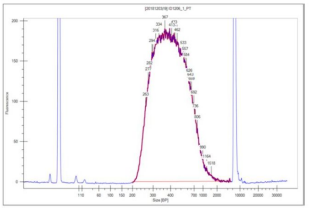


Preparing RNA-seq library with Illumina technology

Kit Ovation® has been used to prepare RNA-seq libraries and RNA-Seq System V2 (Nugen, San Carlos, CA) to prepare cDNA, and kit Ovation® Ultralow System V4 1–96 to prepare the library starting from 10ng cDNA. Final libraries have been quantified by fluorimeter Qubit 2.0 (Invitrogen, Carlsbad, CA) and quality-tested by system Caliper GX (PerkinElmer, Waltham, MA) for batch A71CB205A and Agilent 2100 Bioanalyzer, DNA High Sensitivity Analysis kit (Agilent technologies, Santa Clara, CA) for batch A71CB256A. Set out below the tracing for the two obtained libraries:

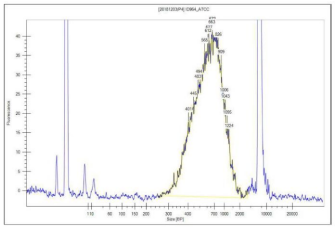


Lotto A71CB205A, libreria RNA-seq diluita 1:40 corsa Caliper GX



Lotto A71CB256A, libreria RNA-seq corsa Bionalyzer

To validate the library preparation workflowuntil the data analysis, Standard ATCC (genomic DNA mix with known composition, 20 Strain Staggered Mix Genomic Material, ATCC® MSA-1003TM) has been used to create a library with kit Ovation® Ultralow System V4 1–96 starting from 10ng DNA. Set out below the tracing from Bioanalyzer of the obtained library:



Sequencing

Libraries have been sequenced on instrument Illumina HiSeq2500 in 'paired-end 125bp' mode, according to Illumina standard instructions. Pipeline Illumina CASAVA version 1.8.2 has been used to process the rough sequences.

Bioinformatic Analysis

Sequences trimming

Adaptors sequences (namely "artificial" oligonucleotides sequences that are introduced during the illumina library preparation) and of DNA bases at low quality read have been removed using **ERNE1** and **Cutadapt2** softwares..

Identification of DNA and cDNA sequences original organisms

Metagenomic analysis has been performed using software **Kraken3** on **database 'Human-Virus-Bacteria_25mer'** (https://ccb.jhu.edu/software/kraken/).

Kraken is a classifier that assigns taxonomic tags to short DNA readings. It tests the k-mers inside a reading and queries a database which includes those k-mers.

Bibliographic references

1. Del Fabbro, C et al. 2013 An extensive evaluation of read trimming effects on Illumina NGS data analysis. Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. PLoS One. 2013 Dec 23;8(12):e85024. doi: 10.1371/journal.pone.0085024. eCollection 2013
2. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, [S.l.], 17 (1): 10-12 (2011). ISSN 2226-6089. Date accessed: 02 Apr. 2015. doi:http://dx.doi.org/10.14806/ej.17.1.200 paper
3. Wood and Salzberg. Kraken: ultrafastmetagenomicsequence classification using exact alignments Genome Biology 2014, 15:R46

CORVELVA
insieme dal 1993

Results of the DNA-seq and RNA-seq analysis carried out on the two batches with the Kraken software

The presence of DNA and RNA is expressed as the number of reads and percentage of reads on the total of the reads produced, attributed by the public databases to the various organisms.

DNA analysis Batch #1 A71CB205A (June 2018)

DNA Seq total reads 3.830.074

Classification		**n° reads**	**% reads**
▪ Homininae (Homo sapiens)		2,853,788	74%
▪ Viruses		538,112	14%
▪ Aves (*Gallus gallus*)		152,256	4%
▪ Proteobacteria (*Bradyrhizobium*) [4]		43,268	1%
▪ Unassigned		193,248	5%

***Viruses classification**		**n° reads**	**% reads**
▪ **dsDNA viruses**		**538,112**	**14%**
	Human alphaherpesvirus 3	537,849	14%
▪ **ssRNA viruses**		**34**	**0,01%**
	Mumps rubulavirus	19	0.0005%
	Measles morbillivirus	12	0.0003%
	Rubella virus	3	0.00008%
▪ **Microviridae** [5]		**131**	**0.003%**
▪ **Retroviridae**		**26**	**0.0007%**
	Avian endogenous retrovirus EAV- HP	7	0.0002%
	Avian erythroblastosis virus	1	0.00003%
	Avian leukosis virus	5	0.0001%
	Human endogenous retrovirus HERV-K(II)	1	0.00003%
	Human endogenous retrovirus K	6	0.0002%
	Human endogenous retrovirus	3	0.00008%
▪ **unclassified bacterial and environmental viruses**		**4**	**0.0001%**

[4] https://it.wikipedia.org/wiki/Proteobacteria
[5] https://en.wikipedia.org/wiki/Microviridae

CORVELVA
insieme dal 1993

RNA analysis Batch #1 A71CB205A

DNA Seq total reads **10.445.038**

Classification		n° reads	% reads
▪ Homininae (Homo sapiens)		9,036,993	87%
▪ Viruses		499,098	5%
▪ Platyhelminthes (Spirometra erinaceieuropaei)		57,805	0,6%
▪ Aves (*Gallus gallus*)		16,361	0,2%
▪ Unassigned		45,660	5%

*Viruses classification		n° reads	% reads
▪ dsDNA viruses		**497,498**	**5%**
	Human alphaherpesvirus 3	497,465	5%
▪ ssRNA viruses		**1.324**	**0,01%**
	Mumps rubulavirus	874	0.008%
	Measles morbillivirus	441	0.004%
	Rubella virus	7	0.00007%
▪ Microviridae		**247**	**0.002%**
▪ Retroviridae		**23**	**0.0002%**
	HERV-H/env60	9	0.0002%
	HERV-H/env62	2	0.00003%
	Human endogenous retrovirus K	6	0.0001%
	Human endogenous retrovirus	2	0.00003%
	Human immunodeficiency virus 1	3	0.0002%
	Alpharetrovirus (Avian viruses)	1	0.00008%
▪ unclassified bacterial and environmental viruses		**5**	**0.0001%**
▪ Hepatitis B virus		**1**	**0.00001%**

CORVELVA
insieme dal 1993

DNA analysis Batch #2 A71CB256A

DNA Seq total reads **5.836.297**

Classification		n° reads	% reads
▪ Homininae (Homo sapiens)		5,150,674	88%
▪ Viruses		643,575	11%
▪ Unassigned		29,634	0,5%

*Viruses classification		n° reads	% reads
▪ dsDNA viruses		**643,549**	**11%**
	Human alphaherpesvirus 3	643,542	11%
▪ Retroviridae		**19**	**0.0003%**
	Avian endogenous retrovirus EAV-HP	13	0.0002%
	Mus musculus mobilized endogenous polytropic provirus	2	0.00003%
	Murine type C retrovirus	1	0.00002%
	Alpharetrovirus	1	0.00002%
▪ Bullavirinae		**5**	**0.00009%**
▪ Saccharomyces 20S RNA narnavirus		**1**	**0.00002%**
▪ Saccharomyces cerevisiae virus L-BC (La)		**1**	**0.00002%**

RNA analysis Batch #2 A71CB256A

DNA Seq total reads **6.171.266**

Classification		n° reads	% reads
▪ Homininae (Homo sapiens)		4,210,032	68%
▪ Proteobacteria		336,053	5.5%
▪ Nematoda (Elaeophora elaphi)		92,290	1.5%
▪ Viruses		419,863	7%
▪ Unassigned		389,837	6%

*Viruses classification		n° reads	% reads
▪ dsDNA viruses		**418,104**	**7%**
	Human alphaherpesvirus 3	417,868	7%
▪ ssRNA viruses			0.01%
	Mumps rubulavirus	508	0.008%
	Measles morbillivirus	217	0.004%

	Rubella virus	Non Determ.	Non Determ.
	Influenza A virus	9	0.0001%
	Cupixi mammarenavirus	30	0.0005%
	Pneumoviridae	6	0.0001%
	Jamestown Canyon virus	12	0.0002%
	Hepacivirus C	30	0.0005%
	Kobuvirus	20	0.0003%
	Enterovirus	2	0.00003%
	Porcine reproductive and respiratory syndrome virus	3	0.00005%
	Coronavirinae	5	0.00008%
	Potyvirus	3	0.00005%
▪ Retroviridae		**99**	**0.002%**
	Human immunodeficiency virus 1	36	0.0006%
	Human endogenous retrovirus K	16	0.0003%
	Simian immunodeficiency virus	1	0.00002%
	Equine infectious anemia virus	1	0.00002%
	Lymphoproliferative disease virus	25	0.0004%
	Avian leukosis virus	1	0.00002%
	Rous sarcoma virus	1	0.00002%
	HERV-H/env62	2	0.00003%
	Red clover bacilliform virus	2	0.00003%
▪ unclassified bacterial and environmental viruses		**88**	**0.001%**
▪ dsRNA viruses		**14**	**0.0002%**
Hepatitis B virus		**2**	**0.00003%**
unclassified RNA viruses ShiM-2016		**9**	**0.0002%**
Mollivirus sibericum		**1**	**0.0002%**

***The RNA-seq library was subsequently sequenced at a very high depth (260.434.942 Illumina reads produced), only in order to highlight the presence of the virus of the RUBELLA. 114 reads attributable to the rubella genome were found, corresponding to a percentage of reads equal to 0.00004%.**

CORVELVA
insieme dal 1993

Results of the DNA-seq analysis carried out with the Kraken software on a genomic standard with a known composition (*20 Strain Staggered Mix Genomic Material, ATCC® MSA-1003TM*)

DNA Seq total reads 4.969.245

Classification	n° reads	% reads	% declared by ATCC
▪ Acinetobacter baumannii	10,735	0.2%	0.18%
▪ Actinomyces odontolyticus	2	0.00004%	0.18%
▪ Bacillus cereus	176,327	3.5%	18%
▪ Bacteroides vulgatus	1,088	0.02%	0.02%
▪ Bifidobacterium adolescentis	489	0.01%	0.02%
▪ Clostridium beijerinckii	123,609	2.5%	1.8%
▪ Cutibacterium acnes	6,528	0.13%	0.18%
▪ Deinococcus radiodurans	745	0.02%	0.02%
▪ Enterococcus faecalis	704	0.01%	0.02%
▪ Escherichia coli	929,837	19%	18%
▪ Helicobacter pylori	4,738	0.1%	0.18%
▪ Lactobacillus gasseri	4,491	0.1%	0.18%
▪ Neisseria meningitidis	9,820	0.19%	0.18%
▪ Porphyromonas gingivalis	578,294	12%	18%
▪ Pseudomonas aeruginosa	152,307	3%	1.8%
▪ Rhodobacter sphaeroides	1,135,927	23%	18%
▪ Staphylococcus aureus	72,598	1.5%	1.8%
▪ Staphylococcus epidermidis	634,940	13%	18%
▪ Streptococcus agalactiae	31,622	0.6%	1.8%
▪ Streptococcus mutans	526,420	11%	18%
▪ Unassigned	14,248	0.3%	0%

From the comparison of these two vaccines it is possible to highlight the following criticalities:

The Priorix Tetra is a vaccine with a high amount of extraneous contaminant DNA of which 80% on average is human, therefore coming from the MRC-5 cells; in Batch #1 there is also 4% of DNA coming from embryo chicken cells; the remaining 20% belongs to viruses (retroviruses, infectious and carcinogens viruses, phages) and adventitious microorganisms such as bacteria, and worms; in the Priorix Tetra vaccine the human genomic DNA is high molecular weight (> 10.000bp) and the total sequential coverage of the entire reference human genome (HG- 19).

From the EMA's answer to our question [6] about the limits imposed on residues of foreign genetic material in vaccines, it appears that in fact there aren't limits for each vaccine but only for some, reported in the monographs of the product; **the maximum limit envisaged ranges from 10 pg [7] to 10 ng**, based on the theoretical calculation of the possibility of foreign genomic DNA to cause oncogenic mutations.
It is noteworthy that regulatory authorities do not require that these contaminations be tested in the final product, but only in the initial preparation phase, and that for the attenuated virus vaccines the purification of these contaminations is a critical step.8 EMA has not provided specific studies on the dangers of fetal residual DNA, which allow to assess the risk of these contaminations to human health, so this limit remains arbitrary today.

In Batch #2 of Priorix Tetra, fetal DNA is about 325 times higher than the maximum limit of 10 ng and 325,000 times higher than the minimum limit of 10 pg.

On the question of contaminating human DNA, the World Health Institute in an official 2011 document entitled 'Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks' argues that what is necessary to take into consideration with respect to rDNA (residual cell DNA) in vaccines is:

- **A.** reduction in the amount of contaminating DNA during the manufacturing process;
- **B.** reduction in the size of the contaminating DNA during the manufacturing process;
- **C.** chemical inactivation of the biological activity of DNA occurred during the manufacturing process.

Taking into account the three requests described above, the product is considered by their regulatory organs (NRA) and control laboratories (NLC) to be at an acceptable level of risk regarding the presence of DNA from the cell substrate, based on (a) and/or (b) and/or (c), when the data show that adequate levels of safety have been achieved.

In particular, in the batches of Priorix Tetra vaccine tested to date, point A. does not occur because the quantity is about 140 times higher than that recommended by the FDA (in Briefing Document September 19, 2012 : Vaccines and Related Biological Products Advisory Committee Meeting) and the EMA, i.e. ≤ 10ng per dose; point B) does not occur because the DNA is high molecular weight (most > 10,000 bp, as can easily be verified using a simple agarose gel to control the quality of the DNA extracted from the vaccine), i.e. 50 times greater than the size recommended by the FDA (200bp or less). Finally, in the same vaccine, point C) does not occur because, as it contains attenuated viruses, a possible chemical inactivation of DNA, would also inactivate the viruses.

Analysis of the Genetic Variants

[6] Quesito EMA: https://www.ivancatalano.eu/wp-content/uploads/2018/05/Letter.pdf
[7] 10pg è la quantità indicativa di DNA contenuta in una cellula; ciò significa che nel vaccino è contenuta una quantità di DNA proveniente da ben 325.000 cellule fetali

With the Next Generation Sequencing technology it is possible to reconstruct the entire sequence of viral DNA and RNA genomes and bacterial genomes present in the sample and compare it with the reference genomes present in the public databases. therefore, technology can also allow to monitor in time how and if the sequence of a viral or bacterial genome changes during the production procedure of a vaccine.

The result of the variant calling (single nucleotide and small insertions/deletions) compared to the reference strains available in the public databases (NCBI, National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/) performed on the samples containing live attenuated viruses or bacteria gave the following results:

Sample 1. - Priorix Tetra

1. The measles virus genome contained in the vaccine is identical to the sequence of **Edmonston Swartz strain** deposited in the databases having accession number AF 266291.1.
2. The number of variants detected was in fact equal to 0;
3. The **mumps** virus genome contained in the vaccine showed a single mutation compared to the **Jeryl-Lynn** viral strain present in public databases with accession number AF 338106.1;
4. The **rubella** virus genome has not been detected;
5. The varicella virus genome contained in the vaccine showed four mutations compared to the **Human herpesvirus 3** present in public databases with accession number AB097932.1;

The sequence of viral antigens / genomes is a strictly confidential data that is not provided by the EMA. There are no guidelines that regulate the analysis of genetic mutations and the study of the effects on human health.

The high frequency of genetic mutations in viruses and bacteria, as well as in the DNA of cell lines in culture, is a problem of great importance with regard to safety, as it is not known how any variants found can modify the infectious capacity and the stimulation of the immune system towards autoimmune reactions.

For example, it is suggested that Efsa now requires the genomic characterization of probiotic strains for human / animal use and subsequently the demonstration of the coincidence, over time, of the sequence of the microorganism compared to that declared, while in the case of vaccines, as Vivotif, as many as 154 genetic variants are tolerated compared to that stated in the data sheet and present in public databases as a reference vaccine strain.

The presence of genetic variants in vaccine samples compared to the strains declared can be considered in our opinion a non-compliance of drugs.

Expansion

In the recent prepublished article on F1000research "Do you cov me? Effect of coverage reduction on species identification and genome reconstruction in complex biological matrices by metagenome shotgun high-throughput sequencing" NGS technology has been used to analyze biological matrices of different types, including two batches of measles-mumps-rubella-varicella vaccine PRIORIX TETRA (GlaxoSmithKline SpA), with the aim of demonstrating how, even from low-coverage NGS sequencing (i.e. of a few hundred thousand sequence fragments to 1 million), it is possible to characterize the biological component in a complex matrix. The next-generation sequencing was already used on vaccine samples in the publication "Deep sequencing reveals persistence of cell-associated mumps vaccine virus in chronic encephalitis" to demonstrate the coincidence between the vaccine virus genome of mumps and the virus found in the brain tissue of an 18-month-old child with SCID who died of encephalitis.
In particular, in the pre-published article on F1000research, it is observed that about 80% of the sequences obtained with NGS technology on the two vaccine samples, consists of human DNA, as impurity present in the working process; the total amount of foreign DNA is about 2 micrograms, coming from the human fetal cell line MRC-5 used to grow rubella and varicella viruses]. The metagenomic analysis performed on these two samples highlights the potential presence of human DNA in all vaccines containing viruses grown in human fetal lines, moreover already verified with technology other than NGS also by Dr. Deisher in "Epidemiologic and Molecular Relationship Between Vaccine Manufacture and Autism Spectrum Disorder Prevalence".

Dr. Theresa Deisher's group in the article "Insertional mutagenesis and autoimmunity induced disease caused by human fetal and retroviral residual toxins in vaccines" states that the levels of residual DNA in MPR, varicella and hepatitis A vaccines available in the United States are way beyond the limit set in the current WHO DNA from immortalized cell lines guideline of 10 ng per dose of vaccine. Although the EMA guideline does not provide for maximum limits for the fetal DNA remaining in the vaccines, Prof. Deisher's group has anyway taken as reference the maximum dose of 10 ng as a consequence of the fact that the short fragments of fetal DNA present in the vaccines have the ability to integrate into the host DNA and can lead to mutagenesis and/or genomic instability as well as an autoimmune response. Furthermore, in some varicella and MPR vaccines, the presence of fragments of endogenous human retrovirus K (HERVK) has been found, which can be re-activated and can facilitate the integration of free DNA into the host genome.
As stated by Dr. Deisher, the danger of retroviral fragments and residual human diploid DNA has not yet been studied in the recipients of the vaccine, although the scientific literature clearly demonstrates the high probability of the dangers of autoimmune and/or insertional mutagenesis due to presence of these residues, and this is a risk to human health that undoubtedly requires serious scientific and epidemiological research.

These results are integrated with the new data on the analysis of a new sample of Priorix Tetra and the reanalysis of the previous ones, as a non-conformity has emerged concerning the vaccine antigens, i.e. the dubious presence of the rubella antigen. From the in-depth analysis of the three Batches it was confirmed that the rubella antigen is not present, because the number of copies per sample is totally negligible in order to have an immunostimulatory effect.
This more in-depth investigation has made it possible to detect the presence of adventitious DNA and RNA, i.e. from viruses, bacteria, fungi and helminths in quantities below the limit of detection of the instrument, and therefore in residual quantities. However, it should be emphasized that the average total amount of foreign DNA varies from 1.9 to 3.7 micrograms, of which about 80% comes from human fetal DNA and the remaining 20% from embryo chicken DNA and adventitious genetic material, thus in non-residual cumulative quantity.
It should be noted that in the various Batches a considerable variability of the adventitious contaminants was noted, but overall the following categories of adventitious microorganisms were found:

▪ **Bacteria**	▪ Proteobacteria
▪ **Worm**	▪ Platyhelminthes ▪ Nematoda
▪ **dsDNA viruses**	▪ Virus of Varicella
▪ **ssRNA viruses**	▪ Rubella, mumps and measles virus ▪ Other viruses including influence A, Cupixi mammarenavirus, Pneumoviridae, Jamestown Canyon, Hepacivirus C, Kobuvirus Enterovirus, Porcine reproductive and respiratory syndrome virus
▪ **Retrovirus**	▪ human and avian endogenous retroviruses ▪ avian virus ▪ human immunodeficiency virus and monkey ▪ murine virus ▪ infectious anemia virus of the horse

	▪ lymphoproliferative disease virus ▪ Rous sarcoma virus
▪ **Other viruses**	▪ Human alphaherpesvirus 3 ▪ Epatite b virus ▪ Yeast virus

References

- Qin J, Li R, Raes J, et al.: A human gut microbial gene catalogue established by metagenomic sequencing.Nature. Nature Publishing Group; 2010; 464(7285): 59–65
- Posada-Cespedes S, Seifert D, Beerenwinkel N. Recent advances in inferring viral diversity from high-throughput sequencing data. Virus Res. 2017 Jul 15;239:17-32. doi: 10.1016/j.virusres.2016.09.016. Epub 2016 Sep 28. Review. PubMed PMID: 27693290.
- Cattonaro F, Spadotto A, Radovic S and Marroni F. Do you cov me? Effect of coverage reduction on species identification and genome reconstruction in complex biological matrices by metagenome shotgun high-throughput sequencing [version 1; referees: awaiting peer review]. F1000Research 2018, 7:1767 (https://doi.org/10.12688/f1000research.16804.1)
- Morfopoulou S, Mee ET, Connaughton SM, Brown JR, Gilmour K, Chong WK, Duprex WP, Ferguson D, Hubank M, Hutchinson C, Kaliakatsos M, McQuaid S, Paine S, Plagnol V, Ruis C, Virasami A, Zhan H, Jacques TS, Schepelmann S, Qasim W, Breuer J. Deep sequencing reveals persistence of cell-associated mumps vaccine virus in chronic encephalitis. Acta Neuropathol. 2017 Jan;133(1):139-147. doi: 10.1007/s00401-016-1629-y. Epub 2016 Oct 21. PubMed PMID: 27770235; PubMed Central PMCID: PMC5209397.
- Deisher TA, Doan NV, Koyama K, Bwabye S. Epidemiologic and Molecular Relationship Between Vaccine Manufacture and Autism Spectrum Disorder Prevalence. Issues Law Med. 2015 Spring;30(1):47-70. PubMed PMID: 26103708.
- Jarzyna P, Doan NV, Deisher TA. Insertional mutagenesis and autoimmunity induced disease caused by human fetal and retroviral residual toxins in vaccines. Issues Law Med. 2016 Fall;31(2):221-234. PubMed PMID: 29108182.