Assessment of the potential integration of the DNA plasmid vaccine CLYNAV into the salmon genome

European Food Safety Authority (EFSA),
Ross Houston, Simon Moxon, Fabien Nogué, Nikoletta Papadopoulou, Matthew Ramon and Elisabeth Waigmann

Abstract

The European Commission mandated EFSA to review a new data package provided by the company Elanco, for the possible integration/non-integration of the DNA plasmid vaccine CLYNAV into the genome of Atlantic salmon (*Salmo salar*) and to indicate whether EFSA agrees with the conclusions drawn by Elanco. The vaccine is injected into fish to confer protection against pancreas disease caused by the salmonid alphavirus. The majority of the experimental data provided by the company was for muscle tissue close to the injection site and for gonadal tissue. EFSA considers that the long persistence of DNA plasmid in muscle tissue close to the injection site and the potential heritability of an integration event in gonad cells support the focus of the assessment on both these tissues. The experimental data did not provide scientifically robust evidence for a true integration event. The company overall concluded that the likelihood of integration is negligible, based on considerations in the context of the company's environmental risk assessment, but did not provide a quantitative value for the rate of integration linked to the term 'negligible'. It is therefore not possible to evaluate this statement specifically with regard to integration rates. EFSA notes that knowledge about homologous and non-homologous integration predicts that integration could occur with certain frequency. Therefore, EFSA has constructed worst-case scenarios leading to upper estimates for possible integration rates of the DNA plasmid vaccine into the Atlantic salmon genome. EFSA concludes that, based on the worst-case scenarios described here and taking into account additional factors decreasing the likelihood of integration, the actual integration rate is likely to be orders of magnitude lower than the upper estimated integration rate calculated in the context of the worst-case scenarios. With the available evidence, the actual integration rate cannot be estimated with more precision.

The EFSA Journal is a publication of the European Food Safety Authority, an agency of the European Union.

Summary

On 29 April 2016, the European Commission mandated the European Food Safety Authority (EFSA) to review a new data package provided by the company Elanco, for the possible integration/non-integration of the DNA plasmid vaccine CLYNAV into the genome of farmed Atlantic salmon (*Salmo salar*). The vaccine was developed to be injected prophylactically into the Atlantic salmon to confer protection against pancreas disease caused by the salmonid alphavirus.

EFSA was requested to review the new data package for the possible integration/non-integration of DNA plasmid into the Atlantic salmon genome and to indicate whether EFSA agrees with the conclusions drawn by Elanco. On 1 July 2016, EFSA accepted the mandate.

The assessment of possible integration/non-integration of a DNA plasmid vaccine into the salmon genome was based on an integration assay, including a next-generation sequencing (NGS) study, aiming to determine true integration event(s); on biodistribution data; on a bioinformatics study assessing potential homology between the injected DNA plasmid sequence and the salmon genome; and on some reflections on integration rates from the environmental risk assessment. Additionally, a theoretical integration rate model in conjunction with biodistribution data provided by the company was taken into account.

The majority of the experimental data provided by the company was for muscle tissue close to the injection site and for gonadal tissue. EFSA considers that the long persistence of DNA plasmid in muscle tissue close to the injection site and the potential heritability of an integration event in gonad cells support the focus of the assessment on both these tissues.

In summary, the experimental data did not provide scientifically robust evidence for a true integration event. However, EFSA notes that knowledge about homologous and non-homologous integration predicts that integration could occur with a certain frequency. Therefore, EFSA has constructed worst-case assumptions leading to upper estimates for possible integration rates of the DNA plasmid vaccine into the Atlantic salmon genome.

Under worst-case assumptions, the upper estimated integration rate for gonadal tissue is derived from the limit of detection (LOD) of the biodistribution study (10 DNA plasmid copies/μg genomic DNA (gDNA)). This corresponds to 1 haploid gonadal tissue cell in 31,250 cells carrying 1 integrated DNA plasmid copy. The company also provided a theoretical model applicable to gonad cells. Based on this model, the integration rate was calculated to be 1 DNA plasmid copy/μg gDNA. This corresponds to 1 gonad cell in 312,500 cells carrying 1 integrated DNA plasmid copy. Both these integration rates have been used by the company in their environmental risk assessments.

Similarly, using worst-case assumptions, the upper estimated integration rate for muscle tissue close to the injection site is derived from measured biodistribution data at the latest time point in the study (170.8 DNA plasmid copies/μg gDNA). Since muscle cells are multinucleate cells, the number of integrated plasmid copies per nucleus containing a diploid genome was calculated, rather than the number of integrated DNA plasmid copies per cell. In the worst-case scenario, 1 diploid nucleus in 915 nuclei would have 1 copy of DNA plasmid integrated in the muscle tissue close to the injection site.

EFSA has identified some factors decreasing the likelihood of these worst-case scenarios occurring and took those into account in its conclusion.

The company overall concluded that for both gonadal tissue and muscle tissue close to the injection site, the likelihood of integration is negligible. The conclusion is based on considerations in the context of the company's environmental risk assessment and no quantitative value for the rate of integration was linked to the term 'negligible'. It is therefore not possible for EFSA to evaluate this statement specifically with regard to the integration rates.

EFSA concludes that, based on the worst-case scenarios described here and taking into account additional factors decreasing the likelihood of integration, the actual integration rate is likely to be orders of magnitude lower than the upper estimated integration rate calculated in the context of the worst-case scenarios. With the available evidence, the actual integration rate cannot be estimated with more precision.

Table of contents

1. Introduction

1.1. Background

The DNA plasmid vaccine CLYNAV was developed by the company Elanco to be injected prophylactically in farmed Atlantic salmon (*Salmo salar*) to confer protection against pancreas disease caused by the salmonid alphavirus.

To assess the legal status of farmed Atlantic salmon vaccinated with CLYNAV with regard to the legislation on genetically modified organisms (GMO), the company, upon suggestion of the European Commission, carried out scientific studies on the integration/non-integration of the DNA plasmid into the Atlantic salmon genome. The data supplied in these studies were assessed by the European Food Safety Authority (EFSA), who concluded that the results from these studies were not sufficient to support the company's conclusion of non-integration of the DNA plasmid vaccine in the salmon genome (EFSA, 2013).

On 29 April 2016, the European Commission mandated EFSA to review a new data package provided by the company for the possible integration/non-integration of DNA plasmid into the fish genome and to indicate whether EFSA agrees with the conclusions drawn by the company. On 1 July 2016, EFSA accepted the mandate.

In order to assess the new data package on the integration/non-integration of DNA plasmid into the fish genome, EFSA created an EFSA ad hoc working group covering the expertise needed to assess the various data packages.

1.2. Terms of Reference as provided by the European Commission

EFSA is requested to review the new data package for the possible integration/non-integration of DNA plasmid into the fish genome and to indicate whether EFSA agrees with the conclusions drawn by Elanco.

2. Data and methodologies

2.1. Data

The data package submitted in the frame of this mandate can be divided into an experimental data package and a theoretical integration model. All data packages, including annexes and appendices, were scrutinised and the studies considered relevant to the mandate's questions are listed below:

- A biodistribution study, already received in the frame of the first assessment, was expanded with two additional time points for muscle tissue close to the site of injection and gonadal tissue (Amendment to Final Study Report No. NAHC/RD/01/12 (22 December 2014)) (Jones, 2014).
- An integration study, including a next-generation sequencing (NGS) study (NAH-14-178) (Jones and Tonita, 2015).
- A bioinformatics study assessing potential homology between the injected plasmid sequence and the salmon genome (NAHC/RE/02/15) (Roy, 2015).
- Risk assessment for CLYNAV vaccine use in Atlantic salmon with theoretical considerations on a potential integration rate in gonadal tissues (Hill et al., 2014).
- Environmental risk assessment for the CLYNAV vaccine with reflections on integration rates (Jones et al., 2016).

A detailed description of each of these studies can be found below (Section 2.1.1 and 2.1.2).

2.1.1. Experimental data package

For the different experimental studies, the company injected intramuscularly a DNA plasmid vaccine (███████ size) at a dose that was at least 2- or 10-fold higher (20 μg or 100 μg of DNA plasmid) than the one intended to be injected for commercial vaccination (5.1–9.4 μg of DNA plasmid). In order for the vaccine to be functional, protein has to be produced and thus a fraction of the plasmid has to be taken up into the nucleus of cells for transcription, before translation takes place. Therefore, genomic and plasmid DNA will be in the same cellular compartment for a period of time.

For the experimental studies, the Atlantic salmon had an average starting weight of 9.0 ± 1.4 g (day 0). Three validated quantitative polymerase chain reaction (qPCR) assays were designed to detect different sequences within the DNA plasmid and were used to detect DNA plasmid (irrespective whether extrachromosomal or integrated) in salmon tissue samples.

- The CMV qPCR assay detects 145 bp of the plasmid with a limit of detection (LOD) of 25 copies/μg genomic DNA (gDNA) from either gonadal and muscle tissue.
- The KanR qPCR assay detects 72 bp of the plasmid with a LOD of 10 copies/μg gDNA from muscle samples or 5 copies/μg gDNA from gonadal tissue.
- The ORF qPCR assay detects 92 bp of the plasmid with a LOD of 10 copies/μg gDNA from either gonadal and muscle tissue.

2.1.1.1. Biodistribution study

The biodistribution study (Jones, 2014) investigates the distribution of the DNA plasmid vaccine in the salmon, from the site of injection (muscle) to the different organs. The tissue samples for the different experiments were collected at 1, 7, 14, 21, 36, 60, 91, 135, 331, 759 and 822 days post-vaccination (DPV). The tissues sampled were from the muscle in and around the site of injection (2 g at DPV 1–331 and 4 g at DPV 759–822), gonadal tissue (300 mg), and other non-injection site organs (gut, spleen, kidney, heart). According to the gonadal development of farmed Atlantic salmon, the presence of gonads and their maturation stage was highly variable in the first 23 months of the sampling.

The data package on the time points 1–331 DPV was previously described by EFSA (EFSA, 2013). From this previous package, only the DNA plasmid levels from gonadal and muscle tissue samples 1 and 14 DPV which are used in the EFSA assessment (detailed in Section 3) are listed (Table 1). The description below focuses on the two additional time points, 759 and 822 days DPV, analysed for the presence of the DNA plasmid vaccine in the muscle and gonadal tissue samples. Detection of DNA plasmid was performed with the CMV qPCR assay. The 2× and 10× DNA plasmid-vaccinated groups were analysed in parallel with the 'not vaccinated' group which received 0.9% saline solution via intramuscular injection.

In the 2× DNA plasmid-vaccinated group, the qPCR assay could not detect DNA plasmid copies in gonadal tissue at 14, 759 and 822 DPV, hence the LOD of the qPCR assay, i.e. 10 DNA plasmid copies/μg gDNA is reported. Also in the 10× plasmid-vaccinated group, the qPCR assay could not detect DNA plasmid copies in gonadal tissue at 759 and 822 DPV, hence again the LOD of 10 DNA plasmid copies/μg gDNA is reported. In gonadal tissue isolated from the 2× DNA plasmid treatment group at 1 DPV, the number of DNA plasmid copies detected was 111.75 copies/μg DNA (Table 1).

In the 2× DNA plasmid-vaccinated group, the mean DNA plasmid copy numbers detected in muscle tissue close to the injection site at 1, 7, and 14 DPV are 4.18×10^6, 6.22×10^6, and 1.55×10^6 copies/μg DNA, respectively. In the same group, the mean DNA plasmid copy numbers detected at 759 and 822 DPV are 88 and 102.50 copies/μg DNA, respectively. In the 10× DNA plasmid-vaccinated group, the mean number of DNA plasmid copies detected from muscle tissue close to the injection site sampled at 759 and 822 DPV are 285.20 and 170.80 DNA plasmid copies/μg gDNA, respectively (Table 1).

Table 1: **Biodistribution study:** Mean DNA plasmid levels detected in gonadal or muscle tissue 1, 7, 14, 759 and 822 DPV

Time point of sampling	Biodistribution study (LOD = 10[a])			
	Gonadal tissue DNA plasmid vaccine dose		Muscle tissue DNA plasmid vaccine dose	
	2×	10×	2×	10×
1 DPV	111.75[a]	NA	4.18×10^6	1.46×10^7
7 DPV	LOD	NA	6.22×10^6	1.3×10^7
14 DPV	LOD	NA	1.55×10^6	1.8×10^7
759 DPV	LOD	LOD	88	285.20
822 DPV	LOD	LOD	102.50	170.80

LOD: limit of detection; NA: not assayed.

(a): DNA plasmid copies/μg gDNA.

2.1.1.2. Integration study

The integration study (Jones and Tonita, 2015) submitted in the frame of this mandate reports on two time points, one in gonadal tissue (1 DPV) and one in muscle tissue close to the injection site (14 DPV), of Atlantic salmon injected with $10\times$ DNA plasmid vaccine dose. The 30 gonadal tissue samples were collected from 96 individual fish, while the muscle tissue samples were collected from 29 individual fish. The intended sample weight for DNA extraction was 200 mg for gonadal tissue (sometimes only reached by pooling independent samples) and 400 mg for muscle tissue samples. Both high molecular weight (HMW) and extrachromosomal DNA plasmid were isolated using a column capture technology and subsequently separated on a gel using field inversion gel electrophoresis (FIGE). The presence of DNA plasmid sequences was analysed by qPCR.

In order to determine the clearance rate of extrachromosomal DNA plasmid copies from the HMW DNA fraction, 30 μg naïve salmon genomic DNA was spiked with different amounts of free DNA plasmid copies (pUK-SPDV-poly2#1). Whereas spikes of up to 10^5 DNA plasmid copies resulted in clearance of plasmid from HMW DNA after three FIGE runs, the qPCR assay could still detect between 410 and 548 plasmid copies of an original 10^7 DNA plasmid spike in 30 μg of naïve salmon muscle genomic DNA.

As a control for possible true integration events of DNA plasmid into the salmon genome and to determine the overall assay sensitivity, the company prepared a Chum salmon (*Oncorhynchus keta*) heart cell line that was transfected with transfection-optimised DNA plasmid containing a similar expression cassette as the DNA plasmid vaccine (salmon gDNA with integrated plasmid designated as CHH1-PDNAV gDNA). Both southern blot and PCR showed the presence of the DNA plasmid, but did not directly demonstrate integration of plasmid sequences into the genomic sequence of the cell line. The company calculated an average of 8.87 equivalents of integrated plasmid per diploid salmon genome or approximately 1.4×10^6 copies/μg gDNA, taking into account the weight of an Atlantic salmon diploid genome being 6.4 pg. A quantity of 30 μg of naïve genomic salmon DNA was spiked with CHH1-PDNAV gDNA containing an equivalent of 10,000 or 60,000 integrated plasmid copies. Since only the 60,000 copies spike/30 μg of naïve genomic salmon DNA could be consistently detected in all replicates, the company set the overall experimental LOD to 2,000 copies/μg gDNA.

For muscle tissue samples, HMW and extrachromosomal DNA plasmid underwent four consecutive rounds of FIGE, while for gonadal tissue samples, the HMW and extrachromosomal DNA was separated in three consecutive rounds of FIGE. All DNA fragments > 17 kb were extracted from the gel and subjected to further qPCR analyses (three qPCR assays for gonadal tissue and the open reading frame (ORF) qPCR assay for muscle samples). No PCR inhibition was found after a PCR inhibitor test on the gonadal tissue samples.

In the gonadal tissue samples collected at 1 DPV, no DNA plasmid could be detected with any of the three qPCR assays after three rounds of FIGE (Table 2). However, for the majority of muscle tissue samples collected at 14 DPV, DNA plasmid could still be found in different levels ranging from 25 (limit of quantification (LOQ)) to 20,946 copies/μg gDNA with the ORF qPCR assay after four rounds of FIGE (Table 2). In order to further determine whether the DNA plasmid remaining in the muscle cell DNA after FIGE separation is integrated into the salmon genomic DNA, the company designed a NGS study.

Table 2: **Integration study:** DNA plasmid levels detected in the gonadal or muscle tissue HMW gDNA fraction 1, 14, 759 and 822 DPV

Time point of sampling	Integration study (LOD = 2,000[a])			
	Gonadal tissue DNA plasmid vaccine dose		Muscle tissue DNA plasmid vaccine dose	
	2×	10×	2×	10×
1 DPV	NA	LOD	NA	NA
14 DPV	NA	NA	NA	25–20,946[a]
759 DPV	NA	NA	NA	NA
822 DPV	NA	NA	NA	NA

LOD: limit of detection; NA: not assayed.
(a): DNA plasmid copies/μg gDNA.

2.1.1.3. Next-generation sequencing study

Muscle tissue samples with detectable plasmid levels after four rounds of FIGE were subjected to NGS analysis; samples from fish injected with a $10\times$ vaccine dose at 14 DPV were analysed (Jones and Tonita, 2015, appendix 7). The company used a plasmid sequence capture next generation sequencing method on post-FIGE HMW DNA from 28 muscle samples combined in 10 pools. The NGS libraries were constructed from 1 μg of post-FIGE muscle tissue DNA using the Covaris fragmentation instrument followed by end repair and adaptor ligation and was size selected for 200 ± 50 bp including the adapter. As a positive control, NGS libraries were also constructed from CHH1-PDNAV gDNA (see Section 2.1.1.2). To enrich for DNA plasmid sequences in the pooled samples, the company designed lockdown probes covering the full DNA plasmid vaccine. The hybridisation enrichment strategy was performed at 47°C overnight. The enriched libraries were sequenced on the Ion Proton™ system.

Raw reads were aligned to reference genomes: *Salmo salar* ICSASG v1 (SS), *Salmo salar* repeat database v1.6 (SSR), positive control transfection-optimised DNA plasmid (IC) and the DNA plasmid vaccine (VD). Torren mapper (tmap) version 4.4.7 was used for the mapping procedure which was used to check the success of the sequencing and enrichment strategies.

In order to identify potential chimeric sequences indicative of an integration event, the company split sequencing reads computationally into two subreads. This was done using an anchor sequence of 35, 40, 45, 50 or 55 nucleotides with the remaining sequence being split into an 'overhang' read. If the overhang was shorter than 25 nucleotides, the entire pair was discarded. Pairs were aligned to the appropriate combined reference, either SS and VD or SS and IC (mapping procedure as detailed above). In order to make the procedure more stringent, only unique alignments were reported. In addition, aligned reads were deduplicated using bamUtilities v1.0.9ph3 to remove identical reads which could be a result of potential PCR duplication.

The company classified an integration event as a read where the anchor aligned to the plasmid reference, whilst its overhang aligned to the *Salmo salar* genome. They required at least three different anchor/overhang read pairs (after deduplication) to span the same junction for it to be considered a true integration event. The company reported detection of junction reads in the positive control (CHH1-PDNAV gDNA) but not in any of the six libraries of pooled 14 DPV muscle samples that generated reads. In addition, four additional libraries were sequenced but failed to yield any reads. A unique event was detected at a single anchor length in one 14 DPV muscle pool but the company stated 'that they do not believe this to be a true reflection of an integration event as in all other runs conducted, the integration events were seen with all the anchor lengths used'.

2.1.1.4. Evaluation of potential for homologous recombination between the plasmid and genomic DNA

Evaluation of the potential of the DNA plasmid vaccine sequence for integration into the host genome through homologous recombination was estimated by performing sequence homology searches between the DNA plasmid vaccine sequence and the public NCBI Nucleotide collection (nr/nt), transcriptome shotgun assembly (TSA) and whole-genome shotgun (WGS) contigs databases of *Salmo salar* (taxid:8030) using BLASTn (Roy, 2015). No element of the DNA plasmid vaccine sequence was identified with sufficient length and identity to support homologous recombination (de Vries and Wackernagel, 2002; Monier et al., 2007; Hülter and Wackernagel, 2008; EFSA, 2009; Overballe-Petersen et al., 2013).

2.1.2. Theoretical integration model

A model to estimate the integration rate of DNA plasmid vaccine into gonad cells of salmon is described in the report by Hill et al. (2014). The model is composed of three distinct steps: 1) binding of the DNA plasmid vaccine to the cell membrane of gonad cells, 2) internalisation of bound plasmid into the cell and 3) integration of internalised plasmid into the genomic DNA.

2.1.2.1. Binding of the DNA plasmid vaccine to the cell membrane of gonad cells

For the first step, the classical binding equation for ligand binding to macromolecules with identical independent binding sites is applied (see for example Price and Dwek, 1979; Bisswanger, 2002). In this context, the plasmid is considered the ligand and the gonad cell is considered the macromolecule with multiple binding sites. Plasmid can be either bound to the cell membrane of gonad cells, or free outside the gonad cells.

$$r = \frac{[P]_{bound}}{[B]} = \frac{n\,[P]}{K_D + [P]}$$

r is defined as the concentration of bound plasmid $[P]_{bound}$ divided by the concentration of the gonad cells $[B]$. The equation establishes a relationship between the concentration of bound plasmid $[P]_{bound}$ and the concentration of free plasmid $[P]$, provided that the dissociation constant K_D of the binding reaction, the number of plasmid binding sites per gonad cell n, and the concentration of gonad cells $[B]$, are known.

The equation allows calculating the fraction of bound plasmid from measurements of total plasmid $[P]_0$ which is the sum of bound and free plasmid.

$$[P]_0 = [P]_{bound} + [P]$$

The following considerations were used to estimate the parameters of the equation:

K_D and n: Since there are no values available for K_D and n that describe the binding of DNA plasmid to gonad cells, data describing λ DNA binding to human white blood cells (Bennett et al., 1985) were used.

[B]: The concentration of gonad cells, $[B]$, was calculated based on data provided in the technical dossier where it was determined that 1 mg of gonad tissue contains 1.5 μg genomic DNA. Furthermore, an estimated molecular mass of the haploid salmon genome of 3.27 pg, corresponding to approx. 3×10^9 bp, was used (Davidson et al., 2010).

$[P]_0$: The concentration of total plasmid outside the gonad cells was assumed to be equal to the concentration of plasmid measured in gonadal tissue at day 1 following injection of a $2\times$ DNA plasmid vaccine dose into muscle tissue (111.75 copies/μg DNA; (Jones, 2014)). This assumption was based on the consideration that in gonadal tissue measurable DNA plasmid concentration was detected only on day 1 after the injection, whereas at later time points no plasmid was detected. Consequently, plasmids measured at day 1 are likely outside the cell, since any plasmid integrated into the genomic DNA would also be measurable at later time points.

Using the parameters estimated as described above, the calculation yielded the result that 77% of plasmids measured in gonad tissue at day 1 are bound to gonad cells.

2.1.2.2. Internalisation of bound plasmid into the cell

For the second step, the assumption was made that all plasmid molecules bound to the cell membrane of gonad cells are internalised into the gonad cells.

2.1.2.3. Integration of internalised plasmid into the genomic DNA

For the third step, the assumption was made that only a fraction of the internalised plasmids is integrated into the genomic DNA of gonad cells. In order to estimate that fraction, data from an integration study by Wang et al. (2004) were used. In this study (Wang et al., 2004), the author injected plasmid DNA into the mouse muscle and, to enhance uptake of plasmid DNA, injection was followed by electroporation in some treatments. Hill et al. (2014) used data derived from treatments combining injection with electroporation since in this case the barrier of internalisation of plasmid DNA into cells can be considered reduced as compared to injection only. The amounts of plasmid DNA measured in muscle tissue at 1 week (first time point measured) and 16 weeks (last time point measured) after injection/electroporation of 50 μg of plasmid DNA were used. Plasmid measured 1 week after injection/electroporation was assumed to be internalised into muscle cells since otherwise it would have dispersed from the electroporation site. Plasmid measured 16 weeks after injection/electroporation was assumed to be integrated into genomic DNA as a worst-case scenario. Applying these measurement values, the fraction of integrated plasmid DNA is estimated to be 0.0117 of internalised plasmid DNA.

2.1.2.4. Calculated integration rate into gonad tissue based on the model by Hill et al. (2014)

Applying steps 1–3 led to a calculated integration rate of 3.2×10^{-6} copies of integrated DNA plasmid vaccine/gonad cell. In other words, 1 gonad cell in approximately 312,500 gonad cells would have 1 copy of DNA plasmid vaccine integrated.

2.1.2.5. Environmental risk assessment for the CLYNAV vaccine with reflections on integration rates (Jones et al., 2016)

While Jones et al., 2016 focuses primarily on environmental risk assessment, it also contains some considerations on integration rates. More specifically, based on the outcome of the biodistribution study in gonadal tissue, where it was shown that at 7 DPV and later time points no plasmid could be detected anymore (see Table 1), the LOD of the biodistribution study (10 copies/μg gDNA) was used as a worst-case scenario for integration in gonad cells.

The company presented also an overall assessment of the potential for integration in gonadal tissue and for muscle tissue close to the injection site. For gonads, the company recalled that published literature to date as well as the studies performed by the company, provide no experimental evidence to support integration in reproductive tissue of DNA plasmid-vaccinated fish or mammals. Furthermore, the company stated that the DNA plasmid vaccine does not persist in gonadal tissue after 1 DPV (Table 1) and that gonad maturation occurs well after the time when plasmid can be detected in gonadal tissue. Overall, the company estimated the likelihood of plasmid integration into gonads as 'so remote to be considered negligible'. For muscle tissue close to the injection site, the company recalled that there is only one published case definitively demonstrating integration of DNA plasmid upon intramuscular injection followed by electroporation into mice muscle (Wang et al., 2004), and that the typical electrophoresis/qPCR based methods cannot clearly distinguish between integrated and non-integrated plasmid. Therefore, the company estimated that the DNA plasmid copies detected in the experimental studies (see Tables 1 and 2) are considered to be an overestimation of the probable level of integration, while the true integration frequency must be extremely low. Overall, the company estimated the likelihood of DNA plasmid integration into injection site tissues as 'negligible'.

2.2. Methodologies

The assessment of possible integration/non-integration of a DNA plasmid vaccine into the salmon genome was based on an integration assay, including a NGS study, aiming to determine true integration event(s); on biodistribution data; on a bioinformatics study assessing potential homology between the injected DNA plasmid sequence and the salmon genome; and on some reflections on integration rates from the environmental risk assessment. Additionally, a theoretical integration rate model in conjunction with biodistribution data provided by the company was taken into account.

3. Assessment

Since the majority of the experimental data provided by the company was for muscle tissue close to the injection site and for gonadal tissue, the assessment focuses on these two tissues. In addition, the potential heritability of an integration event in gonad cells and the long persistence of DNA plasmid in muscle tissue close to the injection site support the focus of the assessment on both these tissues.

It is noted that in animal cells, including fish cells, integration of the DNA into the chromosome of the host may not always follow from nuclear internalisation, and long-term persistence of non-integrated DNA may occur within the nucleus and within cytoplasmic pseudonuclei (Etkin and Pearman, 1987).

For the assessment, the following general considerations apply:

- For the detection of possible integration events in these tissues, qPCR analysis was used (three qPCR assays for gonadal tissue and the ORF qPCR assay for muscle samples). The qPCR assays cover approximately 300 bp of the ▮▮▮▮▮ DNA plasmid sequences. Therefore, only a limited percentage of the DNA plasmid sequences that could potentially be integrated into the salmon genome are covered by the detection methodology. On the other hand, a sequence alignment approach did not detect sufficient sequence homology that might have supported homologous recombination between the DNA plasmid and the salmon genome, and therefore no targeted selection of primers could have been made.
- Biodistribution data aims to detect all DNA plasmid vaccine copies and does not distinguish between integrated and non-integrated DNA plasmid. However, a reduction in the amount of plasmid over time, as detected in the biodistribution assay, suggests that such plasmid most likely represents free DNA plasmid since any plasmid integrated into the genomic DNA would likely also be measurable at later time points. However, at the latest time point in a study, no further information is available on the potential subsequent decline of the amount of DNA plasmid. Therefore, in a worst-case scenario, all DNA plasmid at the latest time point is considered to be integrated.

- In case experimental studies do not yield a measured value, the LOD of the respective methods will be considered in a worst-case scenario, if appropriate.

3.1. Assessment of data obtained for gonadal tissue

3.1.1. Assessment of experimental data

For fish injected with a $10\times$ DNA plasmid vaccine dose, the integration study performed for the gonadal tissue at 1 DPV did not detect an integration event after three FIGE rounds with three qPCR assays (Table 2). However, the LOD for the integration experiment was determined at 2,000 DNA plasmid copies/μg gDNA, meaning that less than 2,000 DNA plasmid copies/μg gDNA would not be detected. Since the sensitivity of the integration study is low as compared to the biodistribution study, the biodistribution data available for this treatment group at the later time points are further considered in the assessment. At 759 and 822 DPV, the biodistribution study did not detect measurable levels of DNA plasmid copies, considering an LOD of 10 DNA plasmid copies/μg gDNA for this study (Table 1).

For fish injected with a $2\times$ DNA plasmid vaccine dose, only data derived from the biodistribution study are available. At 1 DPV, 111.75 DNA plasmid copies/μg gDNA were detected in gonadal tissue (Table 1), whereas at 7 DPV the levels of DNA plasmid copies declined to the LOD (10 copies/μg gDNA). This reduction suggests that the plasmid copy number detected at 1 DPV most likely represents free DNA plasmid since any plasmid integrated into the genomic DNA would also be measurable at later time points. Furthermore, at even later time points (14–822 DPV) the biodistribution study did not detect measurable levels of DNA plasmid copies above the LOD (Table 1).

Therefore, for both the $2\times$ and $10\times$ treatment groups, the latest time point corresponding to the LOD (10 copies/μg gDNA) was considered for calculating an integration rate. In a worst-case scenario, all DNA plasmid copies are assumed to be integrated into the gDNA of salmon gonadal tissue. In this case, the integration rate would be 3.2×10^{-5} copies of integrated DNA plasmid per haploid cell, as calculated by the company (Jones et al., 2016). This corresponds to 1 out of 31,250 gonadal tissue cells with 1 copy of DNA plasmid integrated.

3.1.2. Assessment of the theoretical model

The company has also provided a theoretical model to calculate the integration rate in gonad cells, which was considered a valuable contribution since the above considerations have limitations, i.e. (i) they are primarily based on the biodistribution study, which cannot distinguish between free and integrated DNA plasmid, and (ii) they make use of the LOD rather than actual measurement values. The theoretical model calculates a possible integration rate based on measured levels of DNA plasmid in gonadal tissue at 1 DPV in the biodistribution study, i.e. 111.75 DNA plasmid copies/μg gDNA. Using this model, the company calculated an integration rate of 3.2×10^{-6} copies of integrated DNA plasmid vaccine per gonad cell. Therefore, 1 gonad cell in approximately 312,500 gonad cells would have 1 copy of DNA plasmid vaccine integrated. Based on the assumptions made for the different steps of the model (Section 2.1.2), the results obtained via this model can be considered conservative, i.e. it is likely to provide an overestimated integration rate.

3.2. Assessment of data obtained for muscle tissue close to the injection site

The integration study performed for the muscle tissue close to the injection site at 14 DPV detected between 25 and 20,946 DNA plasmid copies/μg gDNA after four FIGE rounds with the ORF qPCR assay (Table 2). This result should be seen in the context of an assay validation experiment addressing the clearance rate of free plasmid from genomic DNA (described in Section 2.1.1.2). In this experiment, the ORF qPCR assay could still detect between 410 and 548 plasmid copies of an original 10^7 DNA plasmid spike in 30 μg of naïve salmon muscle genomic DNA after three FIGE runs. Considering that at 14 DPV the biodistribution study detects approximately 1.8×10^7 DNA plasmid copies/μg gDNA (5.4×10^8 DNA plasmid copies/30 μg gDNA), it is reasonable to assume that most or all of the DNA plasmid copy levels detected in the integration study (between 25 and 20,946 DNA plasmid copies/μg gDNA) are false positives.

To test for true integration events, a NGS study was performed on the muscle samples analysed in the integration assay. The execution and the data obtained from this experiment had some shortcomings, such as: the hybridisation conditions used for the targeted enrichment (47°C overnight)

were not optimal; the low sequencing coverage; and some samples generated no reads. Furthermore, the company did not provide an LOD for this next generation study. Overall, no conclusions about a possible integration can be drawn from this experiment.

In view of the above considerations, the data from the integration and NGS study provided for muscle cells do not allow a conclusion on possible integration events in the Atlantic salmon gDNA. Therefore, the biodistribution data for this tissue are further considered in the assessment.

For fish injected with a $2\times$ or $10\times$ DNA plasmid vaccine dose, the biodistribution study showed at 1, 7 and 14 DPV high plasmid levels in muscle tissue close to the injection site (ranging from approximately 10^6–10^7 DNA plasmid copies/μg gDNA) which declined to 102.5 and 170.80 DNA plasmid copies/μg gDNA, respectively, at 822 DPV (Table 1). This reduction suggests that most of the plasmid copies detected at early time points most likely represents free DNA plasmid since any plasmid integrated into the genomic DNA would also be measurable at later time points.

However, at the end of the study (822 DPV) for both treatment groups, detectable levels of DNA plasmid vaccine were measured (Table 1). In a worst-case scenario, these levels could be considered to represent plasmid copies integrated into the gDNA of salmon muscle tissue. In this case, the integration rate would be 5.5×10^{-4} ($10\times$ dose) of integrated DNA plasmid copies per genome or 3.3×10^{-4} ($2\times$ dose) of integrated DNA plasmid copies per genome. Muscle cells are multinucleate cells, and it is therefore difficult to derive the number of muscle cells that would contain an integrated copy of DNA plasmid from the above calculated integration rate. However, it is considered possible to express the number of integrated plasmid copies per nucleus containing a diploid genome. Applying this consideration, 1 diploid nucleus in 915 nuclei ($10\times$ dose) or 1 diploid nucleus in 1,524 nuclei ($2\times$ dose) would have 1 copy of DNA plasmid vaccine integrated in the muscle tissue surrounding the injection site.

Since the above considerations are based on a worst-case scenario using the DNA plasmid levels obtained from the biodistribution study which cannot distinguish between free and possibly integrated DNA plasmid, the above calculations are likely to provide an overestimated integration rate.

3.3. Assessment of the potential for integration through homologous recombination

DNA integration can be based on homologous recombination leading to site-specific integration or non-homologous recombination leading to random integration. In animals, the ratio between random integration vs homologous recombination can vary from 1:4 to more than 1,000,000:1 (for a review see Smith, 2001 and literature cited herein). In consequence, the presence of homologous sequences in the DNA plasmid and the salmon genomic DNA could potentially influence the success of integration of the foreign DNA into the salmon chromosome. Moreover, integration through homologous recombination would facilitate detection of stable integration events, as this would permit the selection of primers intended to specifically amplify the border sequences of the integration event. As no element of the DNA plasmid vaccine sequence was identified with sufficient length and identity to support homologous recombination, it can be concluded that in case of integration of the DNA plasmid vaccine into the Atlantic salmon genome this would likely take place via random integration.

4. Conclusions

The majority of the experimental data provided by the company was for muscle tissue close to the injection site and gonadal tissue from Atlantic salmon. Therefore, the assessment focusses on these two tissues. In addition, the long persistence of DNA plasmid in muscle tissue close to the injection site and the potential heritability of an integration event in gonad cells support the focus of the assessment on both these tissues.

In summary, the experimental data did not provide scientifically robust evidence for a true integration event. However, EFSA notes, that knowledge about homologous and non-homologous integration predicts that integration could occur with a certain frequency. Therefore, EFSA has considered the evidence in the frame of the LOD of the methodology applied, based on the information provided by the company (Section 2.1), and has used worst-case assumptions leading to upper estimates for possible integration rates of the DNA plasmid vaccine into the Atlantic salmon genome (Section 3).

Under worst-case assumptions, the upper estimated integration rate for gonadal tissue is derived from the LOD of the biodistribution study (10 DNA plasmid copies/μg gDNA). This corresponds to 1 haploid gonadal tissue cell in 31,250 cells carrying 1 integrated DNA plasmid copy.

The company also provided a theoretical model applicable to gonad cells. Based on this model, the integration rate was calculated to be 1 copy of DNA plasmid/μg gDNA. This corresponds to 1 gonad cell in 312,500 cells carrying 1 integrated DNA plasmid copy. Based on the assumptions made for the different steps of the theoretical model, the results obtained via this model can be considered conservative, i.e. it is likely to provide an overestimated integration rate.

Both these integration rates have been used by the company in their environmental risk assessments.

Similarly, using worst-case assumptions, the upper estimated integration rate for muscle tissue close to the injection site is derived from measured biodistribution data at the latest time point in the study (170.8 DNA plasmid copies/μg gDNA). Muscle cells are multinucleate cells, and it is therefore difficult to derive the number of muscle cells that would contain an integrated copy of DNA plasmid from the above estimated calculated integration rate. However, it is considered possible to express the number of integrated plasmid copies per nucleus containing a diploid genome. Applying this consideration, 1 diploid nucleus in 915 nuclei would have 1 copy of DNA plasmid integrated in the muscle tissue close to the injection site.

Since the above estimations for integration rates are based on worst-case scenarios, some factors decreasing the likelihood of such worst-case scenarios actually occurring are considered below:

- The dose considered in the worst-case scenarios is at least $10\times$ higher than the dose intended to be injected in the salmon in a commercial setting. Therefore, the integration rates calculated in the context of the worst-case scenarios are overestimated.
- The worst-case scenarios use data derived from the biodistribution study which cannot distinguish between extrachromosomal and possibly integrated DNA plasmid, assuming all DNA plasmid to be integrated into the genome. This also contributes to an overestimation of the integration rates as calculated in the worst-case scenarios.
- The theoretical model estimates an integration rate into gonad cells of one order of magnitude less than the corresponding worst-case scenario for gonadal tissue. Since the theoretical model, based on the assumptions made, is likely to provide an overestimated integration rate, this further supports that the worst-case scenario for gonadal tissue is an overestimate.

The company concluded, with respect to the integration study that 'The possibility of experimentally detecting an integration event in vaccinated salmon using the current integration assay is practically nil given that the assay is not sensitive enough to routinely detect an integration event at a level below the actual occurrence rate, which is very likely orders of magnitude below the LOD of the assay'. EFSA is in agreement with this statement but notes that the LOD of the integration study (2,000 DNA plasmid copies/μg gDNA) is significantly higher than the values considered in the worst-case scenarios.

In addition, the company overall concluded that for both gonadal tissue and muscle tissue close to the injection site, the likelihood of integration is negligible (Section 2.1.2.5). The conclusion is based on considerations in the context of the company's environmental risk assessment and no quantitative value on the rate of integration was linked to the term 'negligible'. It is therefore not possible to evaluate this statement specifically with regard to the integration rates.

EFSA concludes that, based on the worst-case scenarios described here and taking into account additional factors decreasing the likelihood of integration, the actual integration rate is likely to be orders of magnitude lower than the upper estimated integration rate calculated in the context of the worst-case scenarios. With the available evidence, the actual integration rate cannot be estimated with more precision.

Documentation provided to EFSA

1) Letter from the European Commission to EFSA to review additional data packages submitted to EC as a follow-up to the 2013 mandate on the DNA plasmid vaccine (Ref. Ares (2013) 104417); received 29 April 2016 (EFSA Incoming No. 139308).
2) Acknowledgement letter from EFSA to EC in response to the mandate letter from EC dated 29 April 2016 (Ref. Ares (2013) 104417); received 20 May 2016.
3) Letter from EMA to EC with list of additional data packages; received 17 June 2016.
4) Letter from the European Commission to EFSA with additional clarifications on the mandate, in response to EFSA letter dated 20 May 2016; received 7 July 2016 (Incoming EFSA No. 141559).
5) Letter from EFSA to the European Commission with update of activities between EMA and EFSA and a request for extension of deadline, in response to letter from EC dated 6 June 2016 (Ref. Ares (2016) 2609869); received 1 July 2016 (EFSA Incoming No 141559).

6) Letter from the European Commission to EFSA, acceptance of deadline extension in response to EFSA's letter dated 1 July 2016; received 14 July 2016 (EFSA Incoming No 144037).
7) Letter from EMA to EFSA, for the CVMP nomination of a Hearing Expert for the EFSA expert working group; received 14 July 2016.

References

Bennett RM, Gabor GT and Merritt MM, 1985. DNA binding to human leukocytes. Journal of Clinical Investigation, 76, 2182–2190.

Bisswanger H, 2002. Macromolecules with identical independent binding sites. In: Bisswanger H (ed.). *Enzyme Kinetics Principle and Methods*. Wiley-VCH Verlag GmbH, Weinheim, Germany.

Davidson WS, Koop BF, Jones SJM, Iturra P, Vidal R, Maass A, Jonassen I, Lien S and Omholt SW, 2010. Sequencing the genome of the Atlantic salmon (*Salmo salar*). Genome Biology, 11, 403–409.

de Vries J and Wackernagel W, 2002. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. Proceedings of the National Academy of Sciences of the USA, 99, 2094–2099.

EFSA (European Food Safety Authority), 2009. Consolidated presentation of the joint Scientific Opinion of the GMO and BIOHAZ Panels on the "Use of Antibiotic Resistance Genes as Marker Genes in Genetically Modified Plants" and the Scientific Opinion of the GMO Panel on "Consequences of the Opinion on the Use of Antibiotic Resistance Genes as Marker Genes in Genetically Modified Plants on Previous EFSA Assessments of Individual GM Plants. EFSA Journal 2009;7(6):1108, 8 pp. doi:10.2903/j.efsa.2009.1108

EFSA (European Food Safety Authority), 2013. Scientific advice on the suitability of data for the assessment of DNA integration into the fish genome of a genetically modified DNA plasmid-based veterinary vaccine. EFSA Journal 2013;11(5):3232, 8 pp, doi:10.2903/j.efsa.2013.3232

Etkin LD and Pearman B, 1987. Distribution, expression and germ line transmission of exogenous DNA sequences following microinjection into Xenopus laevis eggs. Development (Cambridge, England), 99, 15–23.

Hill A, Gale P and Kelly L, 2014. Risk Assessment of CLYNAV vaccine use in Atlantic Salmon. Technical dossier, Elanco Animal Health, 2014, with correction made 18th of September 2015.

Hülter N and Wackernagel W, 2008. Double illegitimate recombination events integrate DNA segments through two different mechanisms during natural transformation of *Acinetobacter baylyi*. Molecular Microbiology, 67, 984–995.

Jones G, 2014. Amendment to Final Study Report No. NAHC/RD/01/12. Assessment of the biodistribution and potential for integration of the pUK-SPDV-poly2#1 plasmid DNA vaccine in Atlantic salmon (*Salmo salar*). Technical dossier, Novartis, December 22, 2014.

Jones G and Tonita P, 2015. NAH-14-178, Assessing the potential non-integration of pUK-SPDVpoly2#1 plasmid into the genome of Atlantic salmon (*Salmo salar*) parr vaccinated intramuscularly with a 10x concentrated 0.05 mL dose of CLYNAV in a non-blinded multisite study. Technical dossier, Novartis Animal Health, May 25, 2015.

Jones G, Park A, Tubbs L and Wright T, 2016. CLYNAV, Salmon Pancreas Disease DNA Vaccine, Environmental Risk Assessment. Technical dossier, Elanco, January 13, 2016.

Monier J-M, Bernillon D, Kay E, Faugier A, Rybalka O, Dessaux Y, Simonet P and Vogel TM, 2007. Detection of potential transgenic plant DNA recipients among soil bacteria. Environmental Biosafety Research, 6, 71–83.

Overballe-Petersen S, Harms K, Orlando LA, Mayar JV, Rasmussen S, Dahl TW, Rosing MT, Poole AM, Sicheritz-Ponten T, Brunak S, Inselmann S, de Vries J, Wackernagel W, Pybus OG, Nielsen R, Johnsen PJ, Nielsen KM and Willerslev E, 2013. Bacterial natural transformation by highly fragmented and damaged DNA. Proceedings of the National Academy of Sciences of the USA, 110, 19860–19865.

Price NC and Dwek RA, 1979. Binding of ligands to macromolecules. In: Price NC and Dwek RA (ed.). *Principles and problems in physical chemistry for biochemists*. Oxford University Press, Oxford, UK.

Roy B, 2015. NAHC/RE/02/15, Report on the homology search finding for the sequence PD NAV in the public NCBI Nucleotide collection (nr/nt), Transcriptome shotgun assembly (TSA) and Whole-genome shotgun (wgs) contigs databases of *Salmo salar* (taxid:8030). Technical dossier, Elanco Animal health, March 24, 2015.

Smith K, 2001. Theoretical mechanisms in targeted and random integration of transgene DNA. Reproduction, Nutrition, Development, 41, 465–485.

Wang Z, Troilo PJ, Wang X, Griffiths TG, Pacchione SJ, Barnum AB, Harper LB, Pauley CJ, Niu Z, Denisova L, Follmer TT, Rizzuto G, Ciliberto G, Fattori E, Monica NL, Manam S and Ledwith BJ, 2004. Detection of integration of plasmid DNA into host genomic DNA following intramuscular injection and electroporation. Gene Therapy, 11, 711–721.

Abbreviations

DPV	days post-vaccination
FIGE	field inversion gel electrophoresis
gDNA	genomic DNA
GMO	genetically modified organisms

HMV	high molecular weight
K_D	dissociation constant
LOD	limit of detection
LOQ	limit of quantification
NCBI	National Center for Biotechnology Information
NGS	next-generation sequencing
ORF	open reading frame
[P]	free plasmid
[P]$_{bound}$	bound plasmid
qPCR	quantitative polymerase chain reaction
tmap	Torren mapper
TSA	transcriptome shotgun assembly
WGS	whole-genome shotgun