Experimental design for the development of a multiplex antigen lateral flow immunoassay detecting the Southern African Territory (SAT) serotypes of foot-and-mouth disease virus

Experimental design for the development of a multiplex antigen lateral flow immunoassay detecting the Southern African Territory (SAT) serotypes of Foot-and-mouth disease virus.

Simone Cavalera [a*], Eugenio Alladio [a], Efrem Alessandro Foglia [b], Santina Grazioli [b], Barbara Colitti [c], Sergio Rosati [c], Chiara Nogarol [d], Fabio Di Nardo [a], Thea Serra [a], Valentina Testa [a], Claudio Baggiani [a], Giampietro Maccabiani [b], Emiliana Brocchi [b], Laura Anfossi [a]

[a.] Department of Chemistry, University of Turin, Via P. Giuria 5, Turin, TO, Italy
[b.] Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia-Romagna, National/OIE/FAO, Reference Centre for FMD and SVD, Via A. Bianchi 9, Brescia, BS, Italy
[c.] Department of Veterinary Science, University of Turin, Largo P. Braccini 5, Grugliasco, TO, Italy
[d.] In3Diagnostic, Largo P.Braccini, 2, Grugliasco, TO, Italy

Abstract

Antigenic Lateral Flow immunoassays (LFIAs), rely on the non-competitive sandwich format, including a detection (labelled) antibody and a capture antibody immobilized onto the analytical membrane. When the same antibody is used for the capture and the detection (single epitope immunoassay), the saturation of analyte epitopes by the probe compromises the capture and lowers the sensitivity. Hence, several factors, including the amount of the probe, the antibody-to-label ratio, and the contact time between the probe and the analyte before reaching the capture antibody, must be adjusted. We explored different designs of experiments (full-factorial, optimal, sub-optimal models) to optimize a multiplex sandwich-type LFIA for the diagnosis and serotyping of two Southern African Territory (SAT) serotypes of the Foot-and-mouth disease virus, and to evaluate the reduction of the number of experiments in the development. Both assays employed single epitope sandwich, so most influencing variables on the sensitivity were studied and individuated. We upgraded a previous device increasing the sensitivity by a factor of two and reached a the visual limit of detection of $10^{3.7}$ and $10^{4.0}$(TCID/mL) for SAT 1 and SAT 2, respectively. The positioning of the capture region along the LFIA strip was the most influent variable to increase the detectability. Furthermore, we confirmed that the 13-optimal DoE was the most convenient approach for designing the device.

Introduction

Foot-and-Mouth Diseases (FMD) is an impactful viral disease widespread throughout the world, particularly in Asia, Africa and the Middle East[1, 2]. WOAH/FAO endorsed a Global FMD Control Strategy to detect rapidly and promptly the disease, and to differentiate circulating serotypes, which include O, A, Asia1, SAT 1 and SAT 2 types. Laboratory-based analysis for virus detection in clinical samples includes Virus Isolation (VI), enzyme-linked immunosorbent assay (ELISA) and nucleic acid detection methods, while "pen side tests" exploiting the Lateral Flow Immunoassays (LFIA) technique have been developed as point-of-care testing (POCT) for the rapid, simple and on-field detection of the virus[3]. Viral and bacterial antigens are early infection biomarkers, and their detection allows for an efficient control of the outbreaks and is essential for monitoring and limiting the spread on the disease[4] compared with serological tests. Also, for FMD diagnosis, the identification of the serotype involved in an outbreak is mandatory to accelerate the

confinement of the infection by implementing vaccination campaigns. This considered, in a previous work, we developed two multiplex LFIAs for the detection and serotyping of the FMD virus: one device aimed at identifying O, A and Asia-1 serotypes (*Eurasia* LFIA currently named LFD1), and the other was developed for the serotyping of SAT 1 and SAT 2 serotypes (*Africa* currently named LFD2). Both devices include PAN-FMDV line and enable the detection of all known FMD serotypes[5]. . The format we used was the sandwich-type immunoassay, in which on monoclonal antibody (mAb) was anchored onto the analytical membrane (capture antibody) and another mAb was labelled with gold nanoparticles (detection antibody, mAb_AuNPs) The LFD1 employed the serotype specific monoclonal antibodies as the capture antibodies and a PAN-reactive mAb as the detector, while the LFD2 used serotype specific mAbs and PAN-reactive mAb coated onto the membrane and a pool of the same mAbs as detection antibodies. While almost equivalent analytical performances to those of a reference antigen-ELISA kit were shown by all the diagnostic lines, an exception for the SAT 2 rendered the LFD2 sub-performing for this serotype[6]. In the previous work, we also encountered for the first time an antigen saturating hook effect (asHE), which occurs when sandwich assays are realized in the lateral flow immunoassay platform using the same antibodies as capture and detection ligand (single-epitope sandwich assay), as in the case of the LFD2[5]. The typical "hook effect" is described as a signal decrease due to the simultaneous reaction of an excess target antigens with both immobilized and labelled antibodies, preventing the formation of a sandwich complex, which in turn leads to a loss of signal intensity[7]. Therefore, the traditional hook effect is explained by the excess of the antigen compared to antibodies (**Fig.1**). On the contrary, the asHE occurs when the excess of the detection antibody masks the analyte inhibiting the binding to the capture antibody.

The asHE was then observed in other studies involving single-epitope sandwich assays for the detection of viral antigens[8][9]. The saturation of the antigen seems to be peculiar of the LFIA format, because the sample encounters the detection antibody first, instead of reacting with the capture antibody as happens usually in ELISA. In addition, the LFIA probe is composed of several antibodies linked to the same-coloured nanoparticles whereas tracers used in ELISA typically or ideally show an antibody-to-label ratio of 1:1. We proposed a possible model for the interpretation of the asHE (**Fig.1a-c)**, based on the competition between capturing and detection antibodies for the same epitope. As such, we suggested three factors that could impact on the asHE. The first factor was the time of contact: the more the time the detection antibody spent with the antigen; the higher the probability to saturate antigen epitopes and to inhibit the binding of the antigen to the capture mAb (**Fig.1a**). The time of contact is directly linked to the distance between the region where the sample is applied and the test line (region where the capture antibody is immobilized). For multiplex LFIA development this aspect is especially critical since a choice must be made between test line orders. The second factor depended on the amount of probe: the higher the number of mAb_AuNPs, the higher the probability to saturate the antigen epitopes, again impairing the binding to the test line (**Fig.1b**).

Finally, the third factor considered the impact of the amount of mAb adsorbed on the single AuNPs, (**Fig.1c**). A high number of antibodies adsorbed on the same signal reporter may increase the affinity of the probe towards the antigen so that the antigen saturation was more probable. Provided that different factors should be studied and that their number increased exponentially because we aimed at developing a multiplex LFIA a high number of experiments was needed Therefore, we explored an alternative to the try and error approach, and applied three designs of experiment (DoE) models to optimize the LFIA configuration with the final goal of increasing the diagnostic sensitivity of the LFD2 and, also to identify which of the factors most influenced the sensitivity. To this aim, we investigated the distance of the test line, the mAb-to-AuNP ratio, and the amount of probe and use the maximization of the test line colour as the main figure-of-merit. So, a second aim of the work was to understand if it was possible to limit the number of experiments by the means of reduced DoE designs. Finally, the optimized device was compared with the pristine LFD2, to confirm the effective gain in sensitivity.

Materials and Methods

Chemicals

Au (III) chloride trihydrate (ACS reagent), anti-mouse immunoglobulin G antibody produced in rabbit (7023), casein sodium salt from milk, sucrose, and bovine serum albumin (BSA) were obtained from Merck/Sigma-Aldrich (St. Louis, MO, USA). Tween20 and other chemicals were purchased from VWR International (Milan, Italy). Nitrocellulose membranes (CNPC-SS12) with cellulose adsorbent pad and glass fibre FR-1 sample pads were purchased by MDI membrane technologies (Ambala, India). Glass fibre conjugate pads were obtained from Merck Millipore (Billerica, MA, USA). Statistical calculations were carried out with Chemometric Agile Tool (CAT) free software. Three anti-FMDV specific mAbs (#2H6 anti-FMDV SAT 2, #HD7 anti-FMDV SAT 1 and #1F10 PAN-FMDV) used both as capture mAbs and mAb_AuNP conjugates were provided by the Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia-Romagna.

Synthesis of mAb_AuNP conjugates

AuNPs with a localized surface plasmon resonance (LSPR) band centred at 525.5 nm and mean diameter of ca. 32 nm were prepared by Au (III) chloride trihydrate reduction with sodium citrate (further detail on the synthesis of AuNPs is reported in the SI)[10, 11]. Mab_AuNPs probes used for the LFIA device development were prepared as described in the previous work with minor modifications for the experimental design and final device. The three levels of amounts of antibody for each mL AuNP optical density 1 (mAb-to-AuNPs) were determined based on a salt-induced aggregation test (further detail on the salt-induced aggregation test is reported in the SI)[12]. Briefly, concentrated sodium chloride was added as the aggregation promoter, to mAb_AuNP conjugates obtained from variable mAb-to-AuNPs. When the mAb_AuNP is sufficiently shielded no salt-induced aggregation occurs. According to the stress test, the stabilizing quantities of mAbs were found to be 4 µg, and 8 µg of #2H6, #1F10 and #HD7 mAb, respectively (**Fig.S1**). Hereafter, the aforementioned amounts are indicated as mAb titre (T). For the experimental design, mAb-to-AuNPs corresponding to T, T/2 and T/4 were used. The Visible spectra of the conjugates were acquired by means of an Agilent Cary 60 spectrophotometer (Agilent, Palo Alto, CA, USA) (**Fig.S2**).

Preparation of the mAb_AuNPs mix for the LFD2.

After the experimental design, the best combinations were defined as T/2 (2ug) #2H6-to-AuNPs with OD2 and T/4 (2ug) #HD7-to-AuNPs with OD3. The #1F10_AuNP was added for the detection of Eurasian serotypes on the test line 3 (T3) and used as previously described[5]. We then produced a mix of the conjugates including the #2H6_AuNP: #HD7_AuNP: #1F10_AuNP in the ratio 2: 2: 1 (OD=5). The mixture was dispensed on the conjugate pad as previously described[5]. Minor interventions on the conjugation procedure were made to better shield the conjugates and limit false positive results due to non-specific interactions. Modifications included using casein in place of BSA in some of the buffers used to prepare the conjugates (details on the casein tests are **reported in the SI**; results are shown in **Table S1**).

Production of the single-test-line LFIA devices for the experimental design and the multiplex LFD2

The various capture antibodies (#2H6, #HD7, #1F10 and rabbit anti-mouse #7023) used for drawing test and control lines of the LFIA devices were diluted in phosphate buffer (20 mM pH 7.4). The concentrations were 1.5mg/mL (#2H6), 1.0mg/mL (#HD7), 1.0mg/mL (#1F10) and 0.5mg/mL (rabbit anti-mouse #7023) and applied at 1 µL/cm onto the nitrocellulose membrane by means of a XYZ3050 platform (Biodot, Irvine, CA,

USA), equipped with BioJetQuanti™ 3000 Line Dispenser for non-contact dispensing. The configurations of the device used in the experimental design and in the multiplex format steps are depicted in **Fig.2**. In the FF-DoE, both for SAT 1 and for SAT 2 detecting systems, the "near" and "far" devices were produced by drawing the serotype specific mAbs in T1 and T2 lines, respectively, distancing about 5mm from each other (**Fig.2a**). After the information obtained from the DoE, the configuration of the multiplex device was: anti-SAT 2 (#2H6) 1.5mg/mL as the T1, anti-SAT 1 (#HD7) 1.0mg/mL as the T2, PAN-FMDV (#1F10) 1.0mg/mL as the T3, and rabbit anti-mouse (7023) 0.3mg/mL as the control lines. that the new device differed from the original one for the reverse disposition of the SAT 1 and SAT 2 test lines (**Fig.2b-c**). Strips were composed by overlapping sample pad, conjugate pad, membrane, and adsorbent pad, and were cut in 4 mm-width by means of a CM4000 guillotine (Biodot, Irvine, CA, USA). Finally, strips were inserted into plastic cassettes (Eximio Biotech, China) to fabricate the ready-to-use LFIA device. Cassettes were stored in the dark in plastic bags containing silica at room temperature until use.

Execution of the two LFD2 devices

For standardization, an inactivated virus suspension was used to fortify the running buffer and was used diluted 10-fold. As the negative control the running buffer was used. The formulation of the running buffer was: 26 mM hydrogen carbonate buffer at pH 7.9 supplemented with 1% v/v tween20, 0.25% w/v casein, 0.02% w/v sodium azide. 80uL of the sample was added to the sample well of the device and the mixture of sample and resuspended gold conjugate was left flowing for 15'.

Experimental designs

Given the experimental specifications, the first step in D-optimal design is to create a candidate set of experimental points. The candidate set is a table with one row for each point (experiment) that is eligible for the experimental design (i.e., a full factorial, in this case). For our study, the original candidate set is a full factorial design for 3 factors with 2 and 3 levels, containing 2*3*3 = 18 possible experiments. The D-optimal design tries to maximize an efficiency parameter, which represents variance of the parameter estimates of the model. The efficiency of the full-factorial DoE is 100% and it depends on the number of points in the experimental design, the number of independent variables in the model, and, if available, the maximum standard error for the prediction over the experimental design points. Therefore, the D-optimal approach aims to define a set of candidate experiments that maximize the efficiency of the DoE, even by reducing the number of experiments. As it is reported in **Fig.S3**, the D-optimal design suggested the preparation of a 13-experiment optimal design (optimal) as it reported the maximum value of efficiency in terms of log (Normalized Determinant). Moreover, a second 9-experiment design(sub-optimal) was performed, representing the minimum number of experiments to be performed in order to obtain consistent evaluations when interpreting the DoE results.

The use of the D-optimal function of CAT software allowed us to obtain the list of the experiments and the experimental conditions of the parameters to be employed when building the different experimental schemes (as reported in **Table 1**). Once the 3 different DoE models have been developed, the p-values of the regression coefficients of the developed multiple regression analysis (MLR) model were evaluated to identify the statistically significant parameters and visualize the nature of the relationships of the parameters involved in the DoE. Then, the coefficient of the variable "distance" was evaluated with respect to the maximum signal to decide which of the SAT sandwiches should be placed in the near (T1) or far (T2) position. Finally, once the selected position was determined, the response surfaces of the DoE models were elaborated to find the combinations with maximum intensity (mAb-to-AuNPs; OD).

Experiments involved in the full-factorial, optimal, and sub-optimal designs.

To furtherly describe and characterize the behaviour of the single-epitope sandwiches, we explored reasonable levels of the variables that are supposed to be the most influencing: distance, optical density, and amount of mAb-to-AuNPs (**Table 1**). The "near" and "far" distance variables were codified as near (T1) =-1 and far (T2) =+1, considering we could draw the test line only in position 1 or 2 in the LFD2. Three levels of amount of mAb-to-AuNPs, defined basing on the salt-induced aggregation test, where T is the minimal stabilizing amount, (T=+1, T/2=0, and T/4=-1) were explored.

Table 1: Design of experiments codifications for each of the two mAbs and experiment involved in the different designs

#exp	Distance[a]	mAb-to-AuNP[b]	Optical density[c]	Full factorial	Optimal	Sub-optimal
1	-1	-1	-1	X	X	X
2	-1	-1	0	X	-	-
3	-1	-1	+1	X	X	-
4	-1	0	-1	X	-	-
5	-1	0	0	X	X	-
6	-1	0	+1	X	-	X
7	-1	+1	-1	X	X	X
8	-1	+1	0	X	-	X
9	-1	+1	+1	X	X	-
10	+1	-1	-1	X	X	X
11	+1	-1	0	X	X	-
12	+1	-1	+1	X	X	X
13	+1	0	-1	X	X	-
14	+1	0	0	X	X	X
15	+1	0	+1	X	X	-
16	+1	+1	-1	X	X	X
17	+1	+1	0	X	X	-
18	+1	+1	+1	X	X	X

[a]distance levels codification: T1=-1, T2=+1
[b]mAb-to-AuNP levels codification: T/4=-1, T/2=0, T=+1
[c]optical density levels codification: 1=-1, 2=0, 3=+1

Similarly, the third variable, the optical density, was explored with three levels (OD3=+1, OD2=0, and OD1=-1). Then, truncated DoE were extracted by means of the D-optimal tool of the CAT software, including 13 for the optimal and 9 for the sub-optimal experiments. The DoE were repeated for each of the two mAbs (anti-SAT 1 and anti-SAT 2) and the experiments were carried in two replicates. As selected factors encompassed both 2 and 3 levels, we needed an approach that could efficiently handle factors with different levels. We opted for a D-optimal design to strike a balance between the number of experiments conducted and the amount of information obtained. D-optimal designs are specifically tailored to handle situations where factors have varying levels, thus allowed us to optimize the allocation of experimental runs, ensuring that we obtained the most information with the fewest experiments. Criteria used to judge the results were: no signal appearing at test lines for the negative control antigen and the more intense colour observed at each line for the specific FMDV type. To compare colour intensity, images of the strips were acquired by a benchtop scanner (OpticSlim 550 scanner, Plustek Technology GmbH, Norderstedt, Germany) and processed by QuantiScan 3.0 software (Biosoft, Cambridge, UK). For the detection of the Eurasian serotypes, also the #1F10 was explored fixing the position (T3), so the half of the experiments were needed.

Statistical analyses were conducted using the software CAT, developed by the Italian Group of Chemometrics (by Prof. Riccardo Leardi *et al.*) of the Italian Society of Chemistry (SCI), freely available on internet, within the R version 3.1.2 version[13][14].

Analytical performance of the LFD2

To investigate the analytical performance of the new LFD2 and to compare it towards the one shown by the former LFD2, the intensity of the test line was quantified as the area of the coloured lines quantified by means of the QuantiScan 3.0 software as described above.

The visual LOD was estimated by serial dilutions of the virus SAT 1 BOT 1/68 and SAT 2 ZIM 5/81; the initial titre of the virus was 10e6.8 $TCID_{50}$/mL for both strains. They were initially 20-fold diluted in the buffer and then 2-fold serial dilutions were prepared and analysed with the two devices (old and new LFD2). The visual LOD was defined as the greater virus dilution at which the colour was still visible at the test line. Colour intensities of the developed strips were quantified from pictures and plotted towards dilution factor.

The reproducibility of the new LFD2 was evaluated as follows: three batches of devices were fabricated on different days. Each batch was used to test the inactivated virus suspension (1:10 in buffer) in four replicates. The coefficient of variation (CV%) of replicates (n=3x4=12) was calculated.

The stability over time was studied by storing some devices at room temperature for 4 months and checking their performance by the application of the inactivated viral suspension as described above. In addition, an accelerated stability study was also conducted, in which devices were stored at 37°C for one week and tested as above.

A series of 9 positive samples, i.e., homogenates of epithelium from infected cattle (**Table S2**) were analysed by the old and new LFD2 and results were visually inspected to confirm the attribution of positivity and the absence of cross-reactivity among lines. The colour intensity of the positive lines was recorded and pairwise compared. Differences were statistically evaluated by applying a pairwise T-test and were considered as significant for $p<0,05$ (95% confidence level). Calculations were carried out by SigmaPlot 12.0 (Inpixon, CA, USA).

Results and Discussion

Results from the FF, 13-optimal, and 9-suboptimal DoEs

Given the experimental specifications, the first step in D-optimal design is to create a candidate set of experimental points. The candidate set is a table with one row for each point (experiment) that is eligible for the experimental design (i.e., a full factorial, in this case). For our study, the original candidate set is a full factorial design for 3 factors with 2 and 3 levels, containing 2*3*3 = 18 possible experiments. The D-optimal design tries to maximize an efficiency parameter, which represents variance of the parameter estimates of the model. The efficiency of the full-factorial DoE is 100% and it depends on the number of points in the experimental design, the number of independent variables in the model, and, if available, the maximum standard error for the prediction over the experimental design points. Therefore, the D-optimal approach aims to define a set of candidate experiments that maximize the efficiency of the DoE, even by reducing the number of experiments. As it is reported in **Fig.S3**, the D-optimal design suggested the preparation of a 13-experiment optimal design (optimal) as it reported the maximum value of efficiency in terms of log (Normalized Determinant). Moreover, a second 9-experiment design (sub-optimal) was performed, representing the minimum number of experiments to be performed in order to obtain consistent evaluations when interpreting the DoE results.

Accordingly, six mAb_AuNPs were synthesized, which were characterised by a red-shift in the LSPR due to the coating with antibodies and passivating proteins[9]. The former AuNPs showed a maximum absorption (λ_{max}) due to the LSPR at 525 nm. After the conjugation process, the #HD7_AuNPs gold conjugates showed a red-shift of the λ_{max} of 6 nm independently on the amounts of mAb added, while the #2H6_AuNPs ones showed a 6, 6.5, and 7 nm red-shift for the T/4, T/2, and T amounts of mAb, respectively. The gold conjugates were inserted into the device in three different concentrations (measured as the optical density of the

Table 2: The ratio between the intensity of the more intense signal from each data set compared to the coefficient of the distance

	Data set	intensity at the top of the surface (a.u.)	coefficient of the distance from the model	saturation effect (%)
#HD7 (anti-SAT 1)	FF	720	-65.2	9
	13-optimal	700	-63.7	9
	9-suboptimal	760	-73.8	10
#2H6 (anti-SAT 2)	FF	200	-21.0	11
	13-optimal	210	-22.3	11
	9-suboptimal	200	-22.2	11

solution used to saturate the conjugate pad). All the experiments were carried by testing 80uL of a 10-fold dilution in the running buffer of the inactivated virus (reference FMDV strains grown in cell cultures) as the positive control. The signals developed were then acquired with the OpticSlim scanner and processed recording the intensity of the colour.

The data were reported in the **Table S3**. The graphs in **Fig.S4-6** show the trends of the increase of the mAb-to-AuNPs ratio on respect to the OD and evidence the decrease of the signal as a function of the optical density for high mAb-to-AuNPs, as a confirmation of the asHE. In addition, as we previously hypothesized, the distance is a parameter that largely affects the asHE, since all the experimental designs showed a coefficient <0 for this parameter (**Fig.S7-8**). By calculating the impact of the saturation, we normalized the coefficient for the maximum level in the response surface (**Fig.S4-6**). The use of the whole set of experiments (FF) allows us understanding that SAT 1 suffered less the impact of the asHE (9%) on respect to SAT 2 (11%). The same conclusion could be extracted from the 13 data set, while the difference was much slighter in the 9-suboptimal data set (10% instead of 11%) (**Table 2**).

Then, considering the sole set in which the SAT 2 assay was carried out as the first test line (T1) and the one in which the SAT 1 assay was carried out as the second test line (T2), we selected the best performing

combination in terms of signal intensity. The description of the response curve (**Fig.S4-6**) suggested that the best approach is using probes with low mAb-to-AuNPs ratio with higher optical density. We hypothesized a behaviour analogous to competitive tests, where the amount of the labelled antibody should be carefully tuned to reach high sensitivity and a compromise between the signal (binding to the competitor antigen) and the binding to the analyte. Moreover, once the sandwich has taken place, the rest of the antibodies on the AuNP surface is probably useless and does not participate to generating the signal through other binding events.

In the different DoE investigated, we found a very diverse description of the experimental space: the FF design and the 13-experiment D-optimal DoE allowed us to compute the quadratic and the interaction terms for the factors under evaluation. On the other hand, the use of the 9-experiment design did not allow us to manage a proper number of degrees of freedom (considering the quadratic and the interaction terms). Therefore, in order to obtain diagnostic plots and compute the variance explained also for this model, no interactions and no quadratic terms were considered for the 9-suboptimal design. A decreasing amount of explained variance was justified by the different models for both the SAT 1 and SAT 2 detecting systems. The results in terms of Explained Variance in Cross-Validation (CV) are reported in the in **Table S4**The FF and the 13-experiment D-optimal designs were capable of describing a larger amount of variance, thus indicating that they can model and interpret the variability of the collected data in a robust way, especially when compared with the 9-experiment models, which did not include the quadratic and the interaction terms in their computation. Moreover, the amount of CV Explained Variance (%) described by SAT 1 system is, on average, higher than the one expressed by the DoE designs of SAT 2 system. This result might suggest that the whole variability of SAT 2 system is not entirely managed by the DoE models as the collected results showed a higher number of random effects. Finally, the evaluation of the response surfaces, explained variance and coefficients of the MLR models allowed us to observe that lowering excessively the number of experiments increased the risk of misleading the configuration so, following the indication from the FF DoE and the 13-optimal DoE, SAT 1 was assigned to the T2 line in the multiplex configuration (**Fig.2c**).

Gold conjugate mix including #HD7_AuNP, #2H6_AuNP, and #1F10_AuNP.

The gold conjugate mix including the three labelled mAbs at the top of their performing conditions, as established by the above-described DoE, were incorporated into the LFIA device and the reactive zones were aligned as follows: SAT 2, SAT 1, and PAN-selective assays. T1 and T2 were decided upon DoE described in the previous section. The DoE to optimize the PAN-selective gold conjugate was carried out by evaluating it in the sole T3 position. The top performing #2H6_AuNP (anti-SAT 2) for the T1 position, was characterised by a mAb-to-AuNPs of 2µg (T/2) per mL of AuNP (OD 1) and an optical density of 2, while for #HD7_AuNP (anti-SAT 1), detected in the T2 position, the optimal was recorded for 4µg (T/2) per mL AuNP (OD 1)) #HD7-to-AuNPs and OD equal to 2. The PAN reactive probe, #1F10_AuNP, was characterised by a mAb-to-AuNPs of 4µg (T) per mL AuNP (OD 1) and an optical density of 1. We then mixed up the three gold conjugates #2H6_AuNP + #HD7_AuNP + #1F10_AuNP in the ratio 2 + 2 + 1. Adjustments of the conjugation procedure were made to better shield the gold nanoparticles and to prevent undesired signals on the negative control, including using casein in place of BSA in some of the buffers (details on studies utilizing casein are **reported in the SI** and results are shown in **Table S1**). The procedure that allowed for completely avoid the false positivity while maintaining a sufficiently appreciable specific signal was the inclusion of 1mg/mL of casein in the second washing, in the dilution, and in the pre-saturation, buffers used in the conjugation process. The use of casein in other conjugation steps, such as, for example, in the overcoating cleaned up the non-specific binding, as well, but significantly reduced the signal of the positive control, therefore, was discarded.

Comparison of the performance of the new LFD2 towards the old one

The optimised high-sensitive LFD2 (new *LFD2*) was employed for testing serial dilution of reference strains of FMDV SAT 1 and FMDV SAT 2 cultured on susceptible cell lines, namely SAT 1/III (WZ)/ZIM/68 and SAT 2/II/ZIM/81 both titrated as $TCID_{50}/mL=10^{6.8}$. The visual LOD (vLOD) was established as the highest dilution which was undoubtedly attributed as positive by observing the strip by the unaided eye by three independent operators (**Figure 4**). The old LFD2 was characterised by a visual LOD of $10^{4.0}$ $TCID_{50}$/mL and $10^{4.3}$ $TCID_{50}$/mL for FMDV SAT 1 and FMDV SAT 2 detection, respectively, while the new LFD2 showed lower vLOD for both the serotypes ($10^{3.7}$ $TCID_{50}$/mL and $10^{4.0}$ $TCID_{50}$/mL for serotypes SAT 1 and SAT 2, respectively). Accordingly, the improvement reached by the optimization process was confirmed.

To investigate the reproducibility of the new LFD2, three independent batches were fabricated and used to analyse the inactivated viral suspension (positive control). The CV% of replicate measurements (n=4) within batch and between batches (n=12) were both below 20%.
Stability over time of the LFD2 was verified by testing the same samples after storing devices for 4 months at room temperature (mean signal variation compared to t=0 was 30%). Moreover, some devices were also stored for 7 days at 37°C, which is considered as an accelerated predictive stability assessment[15]. Also in this study, the colour intensities of the test lines decreased less than 30% (**Table S5**). These values were considered acceptable and promising for the future validation of the new optimised LFD2.

The new LFD2 was compared with the previous device also by testing nine tissue homogenates from three cattle infected with FMDV SAT 1 and six with FMDV SAT 2 (**Fig.4**). All samples were positive to the PAN FMD real-time RT-PCR (Ct value ranging from 11.66 to 19.48) and reference Antigen-ELISA (OD value ranging from 0.97 to 2.89)[16].
The results were visually evaluated, and the quantitative differences were estimated as colour intensity measured at the test lines. The improvement of the detectability was generally confirmed. Three samples showed statistically significant increment of the signal for the new LFD2 compared to the former one. One sample (SAT1_#1) instead showed a lower signal. It should be noted that the new device was designed with the inversion of the order of type-specific lines. As discussed above, the increasing of the distance between the line and the sample application point decreased the detectability, therefore the slight decrement of the SAT 1-specific detection was expected. The SAT 2_#6 sample showed signals barely detectable by both devices. The result may be explained by the low concentration of the virus, measured also by the reference Antigen-ELISA (**Table S2**). Remaining four samples showed a general increased signalfor the new LFD2, though not statistically significant. The intensity of the signals from the PAN-FMD test line (T3) was also compared (**Fig.S9**). Comparing the new LFD2 with the old LFD2, in the first we observed an increase of colour intensity was shown for FMDV SAT 1-positive samples, while all the FMDV SAT 2 positive samples are confirmed to be negative, in accordance with reference Antigen-ELISA results.

A few examples of LFD have been reported previously for the detection and differentiating of SAT FMD virus (**Table 3)**. The visual detection exploiting gold nanoparticles as colorimetric labels, has been the preferred choice to enable the fabrication of affordable analytical tools. Our previous test showed several advantages, such as the versatility (one device was able to diagnosis the infection and to discriminate SAT 1, SAT 2 serotypes), the diagnostic sensitivity (more than 90.5%) and specificity (no-cross reaction detected), and the rapidity (10’ for completing the assay).

Table 3. An overview on recently reported rapid methods for detecting Southern African Territory (SAT) FMDV serotypes.

Biosensor	Detection (signal reporter)	Serotype detected	Serotype differentiated	LOD (LOG $TCID_{50}$/mL)	Time for completing the assay	Ref.
LFD	Colorimetric (gold nanoparticles)	O, A, Asia 1, C, (most SAT 1 and SAT 3, some SAT 2)	No	n.a.[a]	60'	[17]
LFD	Colorimetric (gold nanoparticles)	SAT 2	SAT 2 specific	n.a.[a]	120'	[18]
LFD	Colorimetric (gold nanoparticles)	SAT 2	SAT 2 specific	SAT 2 3.8	30'	[19]
LFD	Colorimetric (gold nanoparticles)	SAT 1, SAT 3	SAT 1 and SAT 3 (SAT 1 LFD cross-react with SAT 2)	SAT 3.1 SAT 3 1.8	30'	[20]
Multiplex LFD	Colorimetric (gold nanoparticles)	Other serotypes, SAT 1, SAT 2	SAT 1, SAT 2	SAT 1 4.9 SAT 2 5.3	10'	[5] [b]
Multiplex LFD	Colorimetric (gold nanoparticles)	Other serotypes, SAT 1, SAT 2	SAT 1, SAT 2	SAT 1 3.7 SAT 2 4.0	10'	This work

[a] not available

[b] only the parameters of the device for the SAT serotypes detection is reported here

We intervened on the same test device, aiming at keeping the same advantages, but upgrading the device in order to increase the analytical sensitivity. In fact, the new LFD2 was able to correctly classify 8 out of the 9 positive epithelial homogenates (as the previous LFD2), however providing mor intense colouring of the lines, which facilitated the unaided eye evaluation of the result and limit the occurrence of uncertain attribution (and false negativity). At the same time, no false positivity was observed, and the samples were always attributed to the pertinent serotype. A significant improvement of the vLOD was also obtained. This result preliminary also suggests an improved diagnostic sensitivity of the new LFD2, which need to be verified by analysing a larger set of samples.

Conclusions

This work aimed at optimizing a single epitope multiplex LFIA to enhance the sensitivity of a previously reported device for detecting and differentiating two FMDV serotypes (LFD2). The multiplexing complicated further the optimization process, by introducing an additional variable (i.e., the order of alignment of the reactive zones). As such, we applied a DoE approach and compared three different cuts of the DoE, in order to investigate the possibility of limiting the number of experiments, without excluding any relevant variables from the study. The D-optimal cut was able to individuate the most influent variable (line position) and find the top of the response surface as well as the FF-DoE, demonstrating that a reduced number of experiments could be carried out saving time and costs, while maintaining the predictive ability. The optimized device (new LFD2*)* showed improved analytical performance with a decrease of about 1 order of magnitude in the

visual LOD. The increased detectability was further confirmed by revealing SAT 1 and SAT 2-type FMDV in epithelial tissues from infected cattle. Full diagnostic validation of the new LFD2 is ongoing, to also investigate the on-field applicability.

References

1. Poonsuk K, Giménez-Lirola L, Zimmerman JJ (2018) A review of foot-and-mouth disease virus (fmdv) testing in livestock with an emphasis on the use of alternative diagnostic specimens. Anim Heal Res Rev 19:100–112. https://doi.org/10.1017/S1466252318000063

2. Admassu B, Getnet K, Shite A, Mohammed S (2015) Review on Foot and Mouth Disease: Distribution and Economic Significance. Acad J Anim Dis 4:160–169. https://doi.org/10.5829/idosi.ajad.2015.4.3.95165

3. Wong CL, Yong CY, Ong HK, et al (2020) Advances in the Diagnosis of Foot-and-Mouth Disease. Front Vet Sci 7:477. https://doi.org/10.3389/fvets.2020.00477

4. Wu X, Chen Q, Li J, Liu Z (2022) Diagnostic techniques for COVID-19: A mini-review. J Virol Methods 301:114437. https://doi.org/10.1016/J.JVIROMET.2021.114437

5. Cavalera S, Russo A, Foglia EA, et al (2022) Design of multiplexing lateral flow immunoassay for detection and typing of foot-and-mouth disease virus using pan-reactive and serotype-specific monoclonal antibodies: Evidence of a new hook effect. Talanta 240:123–155. https://doi.org/10.1016/J.TALANTA.2021.123155

6. Grazioli S, Ferris NP, Dho G, et al (2020) Development and validation of a simplified serotyping ELISA based on monoclonal antibodies for the diagnosis of foot-and-mouth disease virus serotypes O, A, C and Asia 1. Transbound Emerg Dis 67:3005–3015. https://doi.org/10.1111/TBED.13677

7. Ross GMS, Filippini D, Nielen MWF, Salentijn GIJ (2020) Unraveling the Hook Effect: A Comprehensive Study of High Antigen Concentration Effects in Sandwich Lateral Flow Immunoassays. Anal Chem 92:15587–15595. https://doi.org/10.1021/ACS.ANALCHEM.0C03740

8. Cavalera S, Pezzoni G, Grazioli S, et al (2022) Investigation of the “Antigen Hook Effect” in Lateral Flow Sandwich Immunoassay: The Case of Lumpy Skin Disease Virus Detection. Biosensors 12:739. https://doi.org/10.3390/BIOS12090739

9. Cavalera S, Colitti B, De Mia GM, et al (2023) Development of molecular and antigenic-based rapid tests for the identification of African swine fever virus in different tissues. Talanta 258:. https://doi.org/10.1016/J.TALANTA.2023.124443

10. Turkevich J, Stevenson PC, Hillier J (1951) A study of the nucleation and growth processes in the synthesis of colloidal gold. Discuss Faraday Soc 11:55–75. https://doi.org/10.1039/DF9511100055

11. Kimling J, Maier M, Okenve B, et al (2006) Turkevich method for gold nanoparticle synthesis revisited. J Phys Chem B 110:15700–15707. https://doi.org/10.1021/JP061667W

12. Guo YR, Liu SY, Gui WJ, Zhu GN (2009) Gold immunochromatographic assay for simultaneous detection of carbofuran and triazophos in water samples. Anal Biochem 389:32–39. https://doi.org/10.1016/J.AB.2009.03.020

13. Gruppo di Chemiometria della Società Chimica Italiana (SCI). http://gruppochemiometria.it/index.php/software

14. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2014 Vienna, Austria. http://www.r-project.org/.

15. Anfossi L, Di Nardo F, Profiti M, et al (2018) A versatile and sensitive lateral flow immunoassay for the rapid diagnosis of visceral leishmaniasis. Anal Bioanal Chem 410:4123–4134. https://doi.org/10.1007/s00216-018-1067-x

16. Callahan JD, Brown F, Osorio FA, et al (2002) Use of a portable real-time reverse transcriptase-polymerase chain reaction assay for rapid detection of foot-and-mouth disease virus. J Am Vet Med Assoc 220:1636–1642. https://doi.org/10.2460/JAVMA.2002.220.1636

17. Ferris NP, Nordengrahn A, Hutchings GH, et al (2009) Development and laboratory validation of a lateral flow device for the detection of foot-and-mouth disease virus in clinical samples. J Virol Methods 155:10–17. https://doi.org/https://doi.org/10.1016/j.jviromet.2008.09.009

18. T J, Z LL, W R, et al (2011) Development and validation of a lateral flow immunoassay using colloidal gold for the identification of serotype-specific foot-and-mouth disease virus O, A and Asia 1. J Virol Methods 171:74–80. https://doi.org/https://doi.org/10.1016/j.jviromet.2010.10.002

19. Yang M, Mudabuka B, Quizon K, Nfon C (2019) Generation of monoclonal antibodies against foot-and-mouth disease virus SAT 2 and the development of a lateral flow strip test for virus detection. Transbound Emerg Dis 66:1158–1166. https://doi.org/10.1111/tbed.13076

20. Yang M, Mudabuka B, Dueck C, et al (2021) Development of two rapid lateral flow test strips for detection of foot-and-mouth disease virus SAT 1 and SAT 3. J Virol Methods 291:113967