Supplementary information

Ecology of Canine Parvovirus transmission dynamics in a multi-host, multi-pathogen system

Abdelkader Behdenna*, Tiziana Lembo, Olga Calatayud, Sarah Cleaveland, Jo E. B. Halliday, Craig Packer, Felix Lankester, Katie Hampson, Meggan E. Craft, Anna Czupryna, Andrew P. Dobson, Edward J. Dubovi, Eblate Ernest, Robert Fyumagwa, J. Grant C. Hopcraft, Magai T. Kaare, Christine Mentzel, Imam Mzimbiri, David Sutton, Brian Willett, Daniel T. Haydon, Mafalda Viana*

**Corresponding authors:* mafalda.viana@glasgow.ac.uk, abdelkader.behdenna@glasgow.ac.uk

Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, G12 8QQ, UK

S1. Data

The sampling protocol for the lions and dog datasets were described in Viana et al. 2015 [0]. Below we provide a brief description of this protocol, including differences in sample selection criterion for the dogs.

Lion data

Lion populations in the Serengeti National Park and Ngorongoro Conservation Area have been continuously monitored by the Serengeti Lion Project since 1966. Most lions included in this study have been observed since birth and are recognised from natural markings and whisker-spot patterns [1-3]. Consequently, their date of birth is known. For unknown lions, sampled as part of other interventions (such as snare removal), ages were estimated on the basis of nose coloration [4].

Serum samples for serological testing were opportunistically collected as part of the Serengeti National Park management or research interventions (e.g. fitting/removing radio-collars, snare removals and wound treatment; [5]) led by Tanzania National Parks and Tanzania Wildlife Research Institute and Ngorongoro Conservation Area Authority. To maximise samples and ensure independence of observations, if a lion was sampled multiple times and all samples tested positive for canine parvovirus (CPV) antibodies, only the first sample was included in the analyses. Conversely, if all samples were seronegative, only the last sample was included. Finally, if the first sample was seronegative but the second was seropositive, we included both samples in the analysis but the birth year corresponding to the second sample was considered to be the sample year of the first sample. Table S1 contains the number of samples included in the analyses.

Dog data

Domestic dog populations surrounding the Serengeti National Park have been intensively studied since the early 1990s and serological surveys have been conducted since 1992. Dogs were sampled during central-point and house-to-house vaccination campaigns [6] and, in unvaccinated areas, during randomised household surveys. At the time of sam-

pling, information on the age of each dog sampled is obtained by questioning the owner. Previous longitudinal studies in East Africa demonstrated that owner-reported ages are reliable when evaluated as part of specific research studies [7-8]. The reliability of ages recorded during routine vaccination campaigns is likely to be less certain. The dog sampling protocol is described in [9] and the vaccination coverage is shown in Fig. S1.

To ensure that only data from unvaccinated dogs were analysed, we included only dogs from villages that had never been vaccinated and dogs from vaccinated villages that were i) sampled before the onset of vaccination in a given village; or ii) born after the previous vaccination campaign. Given that dogs in vaccinated areas receive vaccination also against rabies and canine distemper virus (CDV), for dogs older than 1 year and originating from vaccinated areas, sera were also tested for antibodies against rabies and CDV. Dogs that had a negative antibody titre against at least one of the three pathogens were considered unvaccinated and included in the analyses. Very young pups (0 - 3 months) were excluded from the analyses to avoid misclassification of CPV status due to possible maternal antibodies. Following this pre-selection that ensured all dogs were unvaccinated, because CPV vaccine can shed and be acquired by unvaccinated susceptible dogs, for the CPV modelling we further restrict the dog samples to those from which information about village vaccination coverage was available for the year of sampling. Table S1 contains the number of samples available and included in the analyses and Fig. 2 shows the seroprevalence over time in dogs and lions.

Figure S1: Proportion of dogs vaccinated during time in the sampled villages only (blue), in the whole Serengeti National Park region (red) and in the western part of the region (black).

Table S1: Number of serum samples from Ngorongoro and Serengeti lions and unvaccinated domestic dogs sampled in areas surrounding the Serengeti National Park available for canine parvovirus serology. Dog sample sizes are also subdivided depending on whether the dog originated from an unvaccinated or vaccinated village and the sample size used in the CPV state-space modelling.

		Non-vaccinated domestic dogs			
Year	**Lions**	**Total**	**Non-vaccinated villages**	**Vaccinated villages**	**Dogs used in CPV models**
1984	12	-	-	-	-
1985	36	-	-	-	-
1986	7	-	-	-	-
1987	26	-	-	-	-
1988	1	-	-	-	-
1989	12	-	-	-	-
1990	2	-	-	-	-
1991	19	-	-	-	-
1992	4	223	223	-	-
1993	8	155	155	-	-
1994	56	240	240	-	-
1995	-	-	-	-	-
1996	1	181	180	1	178
1997	15	643	458	185	298
1998	28	921	658	263	398
1999	10	486	455	31	-
2000	17	90	23	67	45
2001	21	250	134	116	-
2002	11	151	19	132	-
2003	29	550	508	42	422
2004	32	1073	586	487	350
2005	37	403	35	368	164
2006	29	672	188	484	282
2007	17	307	12	295	26
2008	6	84	-	84	62
2009	4	157	-	157	66
2010	6	133	-	133	99
2011	11	72	-	72	58
2012	1	75	-	75	-
Total	458	6866	3874	2992	2368

Serological assays

All CPV serology was carried out using neutralisations assays. Sera were analysed at Intervet UK, Animal Health Diagnostic Center at Cornell (New York, USA) and University of Glasgow (UK). Protocols were broadly similar in all laboratories, including viral strains used for the assay. We used a cut-off titer value equivalent to a 1:16 dilution to evaluate prior exposure. This value was the minimum dilution consistently used across all samples and is consistent with other studies of CDV exposure in wild carnivore species [9-11]. Sera from older dogs originating from vaccinated areas and with uncertain vaccination status were additionally tested for antibodies against rabies and CPV (see above). Antibody titres for CPV were determined by hemagglutination inhibition (HAI) testing [12] at Intervet (United Kingdom) and the Animal Health Diagnostic Center (Cornell University, Ithaca, NY). Dog CPV serology from samples collected from 2008 to 2012 was carried out using ELISAs (IgG) at the University of Glasgow.

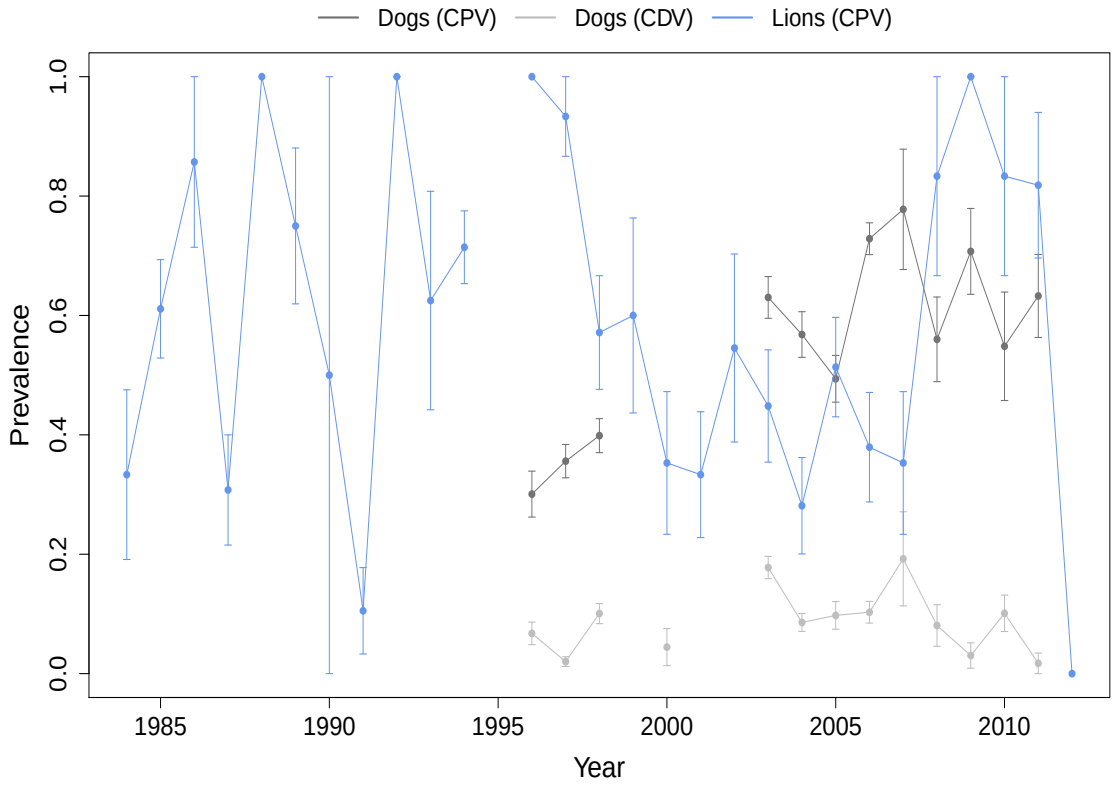

Figure S2: Canine distemper virus (light grey), dogs canine parvovirus (grey) and lions CPV (blue) observed seroprevalence over time, the vertical lines correspond to the standard error.

S2. Serology models

The methodology largely follows from Viana et al 2015 [0]. Briefly we use two Bayesian state-space models. They disentangle the natural from the vaccine acquired CPV immunity but model 1 is used to (i) characterize the within and between-species dynamics of CPV natural infection in dogs and lions; and (ii) quantify the impact of the vaccination programs on those infection dynamics; and model 2 to (iii) investigate the co-infection dynamics between CPV and CDV in domestic dogs in the Serengeti from serology data and vaccination history of each individual. Briefly, the models are all composed of two parts: i) a *Biological process* which characterises the mechanism of infection; and ii) an *Observation process* which confronts the processes underlying the generation of the observed data (mostly at the population level) with the individual level data. Ultimately, the state-space model describes the combination of stochastic processes giving rise to the data in Figure S2.

Biological process

1. Dynamics of CPV at the domestic-wildlife interface and effect of vaccination in dogs

We are interested in estimating, $h_s(t)$, the proportion of infected individuals of species s at year t, where t ranges from 1 to 39 and corresponds to a time series from 1973 (corresponding to the earliest birth date of an individual) to 2012 (last sampling year). Here, this proportion is defined through a logit transformation such that:

$$h_s(t) = \frac{exp(H_s(t))}{1 + exp(H_s(t))} \tag{1}$$

where $H_s(t)$ is the predictor of $h_s(t)$, defined as a stochastic realisation from a Gaussian process:

$$H_s(t) \sim N(\bar{H}_s(t), \sigma_{s,t}) \tag{2}$$

The variability $\sigma_{s,t}$ corresponds to the annual variation around the expectation $\bar{H}_s(t)$. The prior distribution for $\sigma_{s,t}$ is defined in Table S2. $\bar{H}_s(t)$ is formulated as a linear function of covariates that describe the CPV transmission process in domestic dogs and lions. In their most parameterised forms, the expected predictor for dogs and lions were:

$$\bar{H}_{lions,t} = \beta_0 + \beta_1 t + \beta_2 H_{lions,t-1} + \beta_3 H_{lions,t-2} + \beta 4 H_{dogs,t-1} \quad (3)$$

$$\bar{H}_{dogs,t} = \omega_0 + \omega_1 t + \omega_2 H_{dogs,t-1} + \omega_3 H_{dogs,t-2} + \omega_4 H_{lions,t-1} - \omega_5 C_{t-1} \quad (4)$$

The parameters β_0 and ω_0 corresponding to the intercept, which is the baseline probability of infection, while the coefficient β_1 and ω_1 correspond to the coefficient of a linear trend on time. The coefficients β_2, β_3, ω_2 and ω_3 for the dogs and the lions correspond to the coefficients of an autoregressive component (AR) with lag 1 and 2, respectively. The AR terms were used to emulate the disease dynamics as they enable the generation of persistent cycles in the infection dynamics. Given that the questions we ask are at an annual scale, the magnitude of these AR components are thought to be appropriate. The parameters β_4 for the lions and ω_4 for the dogs correspond to the cross-species transmission coefficients , and are defined as an autoregression of lag 1 on the probability of infection of the other species (i.e. β_4 corresponds to the transmission from dogs to lions and ω_4 to the transmission from lions to dogs). The difference between the lion and dog linear predictors is that the dogs might be directly affected by vaccination. Thus the impact of regional vaccination coverage (C_t) on the domestic dog annual CPV seroprevalence was then included as a lagged covariate with coefficient ω_5.. The priors for all these coefficients are defined in Table S2. When needed we constrained the priors to be biologically plausible. For example, the priors for the vaccination parameters were described as strictly positive since vaccination cannot increase infectivity. Owing to the AR lags, priors drawn from a normal distribution with mean ω_0 and variance 10^3 were allocated to the first two time points, $\bar{H}_{t=1,2}$.

At the individual level, the probability of an individual i of species s becoming infected at time t, denoted $u_{i,s}(t)$, is defined as:

$$u_{i,s}(t) = \frac{exp(U_{i,s}(t)}{1 + exp(U_{i,s}(t))} \quad (5)$$

where $U_{i,s}(t)$ is the predictor of $u_{i,s}(t)$ and is modelled as a Gaussian term such that:

$$U_{i,s}(t) \sim N(\bar{U}_{i,s}(t), \psi_{i,s,t}) \quad (6)$$

The stochasticity generated by ψ is the expression of individual variation which increases with the duration of the time interval between the years of birth and sampling; i.e. the uncertainty in the time of exposure of an individual with a greater time interval between birth and sampling will be higher than one with a short interval. The prior distribution for $\psi_{i,s,t}$ is defined in Table S2. $\bar{U}_{i,s}(t)$ is defined as:

$$\bar{U}_{i,lions}(t) = H_{i,lions}(t) + \epsilon_{s,i,t} \tag{7}$$

$$\bar{U}_{i,dogs}(t) = H_{i,dogs}(t) + v_1 V_{i,t} + \epsilon_{s,i,t} \tag{8}$$

The error term ϵ enables additional individual uncertainty for species s, individual i at time t. For dogs, $\bar{U}_{i,dogs}(t)$ also takes into account vaccine shedding and is here that the natural infection can be disentangled from the vaccine acquired "infection". The coefficient v_1 represents the rate governing the impact of the village-level vaccination coverage (V) on the probability of infection of a dog. The prior for this parameter is defined in Table S2.

Table S2: Prior distributions for the parameters used to model the lion and domestic dog populations annual probability of canine parvovirus infection. AR corresponds to the autoregression.

Species	**Variable**	**Parameter**	**Distribution**	**Prior**
Lions	Correct detection	q^+	Beta	$\sim beta(25, 0.5)$
	False detection	q^-	Beta	$\sim beta(0.5, 25)$
	Variance	$\psi\|\sigma$	Normal	$\sim N(0, \tau^{2^{-1}})$
	Standard deviation	τ	Uniform	$\sim U(0, 5)$
	Intercept	β_0	Normal	$\sim N(0, 0.001)$
	Linear trend	β_1	Normal	$\sim N(0, 0.001)$
	AR(1)	β_2	Normal	$\sim N(0, 0.1)$
	AR(2)	β_3	Normal	$\sim N(0, 0.1)$
	Dog-to-lion transmission	β_4	Exponential	$\sim exp(0.5)$
Dogs	Correct detection	q^+	Beta	$\sim beta(25, 0.5)$
	False detection	q^-	Beta	$\sim beta(0.5, 25)$
	Variance	$\psi\|\sigma$	Normal	$\sim N(0, \tau^{2^{-1}})$
	Standard deviation	τ	Uniform	$\sim U(0, 5)$
	Village status (lag 1)	v_1	Exponential	$\sim exp(0.5)$
	Village status (lag 2)	v_2	Exponential	$\sim exp(0.5)$
	Intercept	ω_0	Normal	$\sim N(0, 0.001)$
	Linear trend	ω_1	Normal	$\sim N(0, 0.001)$
	AR(1)	ω_2	Normal	$\sim N(0, 0.1)$
	AR(2)	ω_3	Normal	$\sim N(0, 0.1)$
	Lion-to-dog transmission	ω_4	Exponential	$\sim exp(0.5)$
	Regional vacc. (lag 1)	ω_5	Exponential	$\sim exp(0.5)$
	Shedding parameter	v_1	Beta	$\sim beta(3, 12)$

*Normal distributions are expressed in terms of mean and precision.

3. Co-infection patterns between CPV and CDV in dogs

Finally, we are here interested in estimating, $h_d(t)$, the proportion of dogs infected infected by the disease d at year t, where t ranges from 1 to 27 and corresponds to a time series from 1992 to 2008. The disease d can be either CPV, CDV, or another term called "both", corresponding to the proportion of dogs infected by CPV and CDV at the same time. Again, this proportion is defined through a logit transformation such that:

$$h_d(t) = \frac{exp(H_d(t))}{1 + exp(H_d(t))} \tag{9}$$

where $H_s(t)$ is the predictor of $h_d(t)$, defined as a stochastic realisation from a Gaussian process:

$$H_d(t) \sim N(\bar{H}_d(t), \sigma_{d,t}) \tag{10}$$

In their most parameterised forms, the expected predictor for CPV, CDV and *both* were:

$$\bar{H}_{CPV,t} = \beta_0 + \beta_1 t + \beta_2 H_{CPV,t-1} + \beta_3 H_{CPV,t-2} + \beta 4 H_{CDV,t-1} - \beta 5 C_{t-1} \tag{11}$$

$$\bar{H}_{CDV,t} = \omega_0 + \omega_1 t + \omega_2 H_{CDV,t-1} + \omega_3 H_{CDV,t-2} + \omega_4 H_{CPV,t-1} - \omega_5 C_{t-1} \tag{12}$$

$$\bar{H}_{both,t} = \gamma_0 H_{CPV,t} H_{CDV,t} + \gamma_1 H_{CPV,t-1} H_{CDV,t} + \gamma_2 H_{CPV,t} H_{CDV,t-1} \tag{13}$$

The parameters β_0 and ω_0 corresponding to the intercept, the coefficient of a linear trend on time (β_1 and ω_1), the autoregressive components (β_2, β_3, ω_2 and ω_3) are similar to the ones defined in equations (6), (9) and (10), respectively for CPV and CDV. The parameters β_4 for CPV and ω_4 for CDV correspond to the influence of a disease on the other, and are defined as an autoregression of lag 1 on the probability of infection by the other disease (i.e. β_4 corresponds to the increase of probability of a dog being infected by CPV if it was infected by CDV at t-1 and reciprocally for ω_4). Because we don't see an effect of the lag 2 vaccination terms, we reduced the required computation time to run the model by only considering the vaccination terms with a 1-year lag for CDV and CPV. The parameter γ_0 corresponds to a term of independence: if the two diseases have independent dynamics, then the probability of being infected by both at time t is the product $H_{CPV,t} H_{CDV,t}$. γ_1 corresponds to the effect of CPV infection on the coinfection with a 1-year lag, while $gamma_2$ corresponds to the effect of CDV infection on the coinfection, also with a 1-year

time-lag. The priors for all these coefficients are defined in Table S3.

At the individual level, the probability of an individual i of species s becoming infected at time t, denoted $u_{i,s}(t)$, is defined as:

$$u_{i,s}(t) = \frac{exp(U_{i,d}(t)}{1 + exp(U_{i,d}(t))} \tag{14}$$

where $U_{i,d}(t)$ is the predictor of $u_{i,d}(t)$ and is modelled as a Gaussian term such that:

$$U_{i,d}(t) \sim N(\bar{U}_{i,d}(t), \psi_{i,d,t}) \tag{15}$$

The prior distribution for $\psi_{i,d,t}$ is defined in Table S4. $\bar{U}_{i,d}(t)$ is defined as:

$$\bar{U}_{i,CPV}(t) = H_{i,CPV}(t) + v_1 V_{i,t} \tag{16}$$

$$\bar{U}_{i,CDV}(t) = H_{i,CDV}(t) \tag{17}$$

$$\bar{U}_{i,both}(t) = H_{i,both}(t) \tag{18}$$

Table S3: Prior distributions for the parameters used to model the lion and domestic dog populations annual probability of canine distemper virus infection. AR corresponds to the autoregression.

Disease	**Variable**	**Parameter**	**Distribution**	**Prior**
CPV	Correct detection	q^+	Beta	$\sim beta(25, 0.5)$
	False detection	q^-	Beta	$\sim beta(0.5, 25)$
	Variance	$\psi\|\sigma$	Normal	$\sim N(0, \tau^{2^{-1}})$
	Standard deviation	τ	Uniform	$\sim U(0, 5)$
	Intercept	β_0	Normal	$\sim N(0, 0.001)$
	Linear trend	β_1	Normal	$\sim N(0, 0.001)$
	AR(1)	β_2	Normal	$\sim N(0, 0.1)$
	AR(2)	β_3	Normal	$\sim N(0, 0.1)$
	CDV-to-CPV transmission	β_4	Normal	$\sim N(0, 0.1)$
	Regional vacc. (lag 1)	β_5	Normal	$\sim N(0, 0.01)$
	Shedding parameter	v_1	Beta	$\sim beta(3, 12)$
CDV	Correct detection	q^+	Beta	$\sim beta(25, 0.5)$
	False detection	q^-	Beta	$\sim beta(0.5, 25)$
	Variance	$\psi\|\sigma$	Normal	$\sim N(0, \tau^{2^{-1}})$
	Standard deviation	τ	Uniform	$\sim U(0, 5)$
	Intercept	ω_0	Normal	$\sim N(0, 0.001)$
	Linear trend	ω_1	Normal	$\sim N(0, 0.001)$
	AR(1)	ω_2	Normal	$\sim N(0, 0.1)$
	AR(2)	ω_3	Normal	$\sim N(0, 0.1)$
	CPV-to-CDV transmission	ω_4	Normal	$\sim N(0, 0.1)$
	Regional vacc. (lag 1)	ω_5	Normal	$\sim N(0, 0.01)$
Both	Correct detection	q^+	Beta	$\sim beta(25, 0.5)$
	False detection	q^-	Beta	$\sim beta(0.5, 25)$
	Variance	$\psi\|\sigma$	Normal	$\sim N(0, \tau^{2^{-1}})$
	Standard deviation	τ	Uniform	$\sim U(0, 5)$
	Independent infections	γ_0	Normal	$\sim N(0, 0.1)$
	CPV-to-coinfection	γ_1	Normal	$\sim N(0, 0.1)$
	CPV-to-coinfection	γ_2	Normal	$\sim N(0, 0.1)$

*Normal distributions are expressed in terms of mean and precision.

Observation process

In this section, we describe the observation process corresponding to the domestic-wildlife interface model. The observation processes corresponding to the co-infection model are defined the same way.

Assuming that the i^{th} individual of species s was born in year t_i and was sampled in year T_i, the probability $r_{i,s}(T_i)$ that, at sampling, it was in fact seropositive is:

$$r_{i,s}(T_i) = 1 - \prod_{t=t_i}^{T_i}(1 - u_{i,s}(t)) \tag{19}$$

where, $u_{i,s}(t)$ is the probability of individual i, of species s, becoming infected at time t as described above in equations (3-4). This probability links the observation process at the individual level with the biological process at the population level. We estimate the likelihood of being infected between birth and sampling years because it is impossible to identify the exact time of exposure from a serology test. Once an individual is infected and recovers from CDV it gains life-long immunity, hence, after infection, the individual will always test positive in the serology assay. This means that the time-series investigated starts in the year when the first individual from both species was born, i.e. 1970; and ends in the last year for which data are available, i.e. 2012. As a consequence, larger uncertainty should be expected in the initial years of the time-series when no samples, hence no serological data, were available, but the first lions were already born.

In addition to the inherent difficulty in detecting antibodies, test results from serological assays are typically sensitive to cut-off thresholds [13]. This means that the true disease status of an individual does not always correspond to the test result. In order to account for this potential misclassification of the animal disease status, in Table S4 we introduce probabilities of Type I and Type II errors.

The likelihood that an individual i is detected as seropositive ($P(X_i = 1)$) or seronegative ($P(X_i = 0)$) is based on serology data X and was defined in our model as:

$$P(X_i = 1) = r_{i,s}(T_i)q^+ + (1 - r_{i,s})(1 - q^-) \tag{20}$$

$$P(X_i = 0) = r_{i,s}(T_i)(1 - q^+) + (1 - r_{i,s})q^- \tag{21}$$

Table S4: Probabilities (q) associated with canine distemper virus serological misclassification.

		True state	
		+	-
Test	+	q^+	$1-q^-$
result	-	$1-q^+$	q^-

Based on serological literature [13], we expect high correct detection (q) and low false detection ($1-q$). See Table S3 for further details on these priors.

The total likelihood of the data X under the model and parameters was:

$$P(X_{i,s}) = \prod_{i=1}^{n} x_{i,s} P(X_{i,s}=1) + (1-x_{i,s}) P(X_{i,s}=0) \tag{22}$$

where n corresponds to the number of samples taken from each species and $x_{i,s}$ corresponds to individual draws from the data X. The likelihood of the data X from individual i and species s was generated from a Bernoulli distribution with success probability P, i.e. probability of getting a seropositive individual upon testing, such that:

$$X_{i,s} \sim Bernoulli(P(X_{i,s}=1)) \tag{23}$$

Where $X_s = 1$ corresponds to a CDV positive titer and $X_s = 0$ to a CDV negative titer. If both realisations are equally likely, $P(X_s=1) = P(X_s=0) = 0.5$.

Prior sensitivity

We explored the sensitivity of the model results to the prior distributions by constraining and widening the allocated distribution range. The posterior distributions for most parameters and the annual proportion of infected individuals $h_s(t)$ remained similar, suggesting that the parameter estimates are not sensitive to the priors chosen (Tables S2-4). Wider priors for q^+ and q^- generated convergence issues, but (i) given that these were highly constrained to account for the knowledge that there is a low probability of false detection and high probability of correct detection; and (ii) given that the overall $h_s(t)$

pattern remained largely unchanged; we find our results robust. The posterior distributions for the cross-species transmission terms were sensitive to wide priors (e.g. $\sim exp(10)$) but these are not thought to be plausible values for these parameters. These posterior distributions were not sensitive to the small changes in the priors (e.g. mean 0.2 to 2).

Model fit, convergence and sensitivity analysis

Both models were fitted using JAGS software (22) which uses Gibbs sampling to generate posterior distributions of the parameters given the likelihood, prior distributions and the data itself. We ran our models for at least $50 \cdot 10^4$ iterations with burn-in of $25 \cdot 10^4$ to achieve convergence. The models from which we draw inferences are the ones that address our full set of scientific questions and both converged well and generate validated fits. To ensure these models perform well, we evaluate the numerical robustness in the form of convergence and mixing criteria (e.g. visual inspection of the chains, as well as the Gelman-Rubin statistics [14,15] shown in Table S5 and S6) and goodness-of-fit, by, for example, validating the posteriors against the priors and comparing the observed with the estimated values.

We used a forward prediction approach to investigate the (1) effect of regional-level vaccination on the annual proportion of infected domestic dogs, and of the presence of cross-species transmission from (2) lions-to-dogs and (3) dogs-to-lions on their respective annual proportion of infected individuals ($h_s(t)$). To ascertain the predictive power of the model, we compared the estimates from equation 3 (i.e. infection hazard of dogs at time t) with its prediction estimates (Figure S3). These results shows that the median and credible intervals of the estimate (grey) and prediction (red) are very similar. This suggests a good predictive power of at least the mean of the hazard. Consequently, model predictions were generated from Equations 3 and 4 and these were compared to those in which the parameter of interests (e.g. ω_5 for vaccination, β_4 for dog-to-lions and ω_4 for lion-to-dogs transmission) were set to zero.

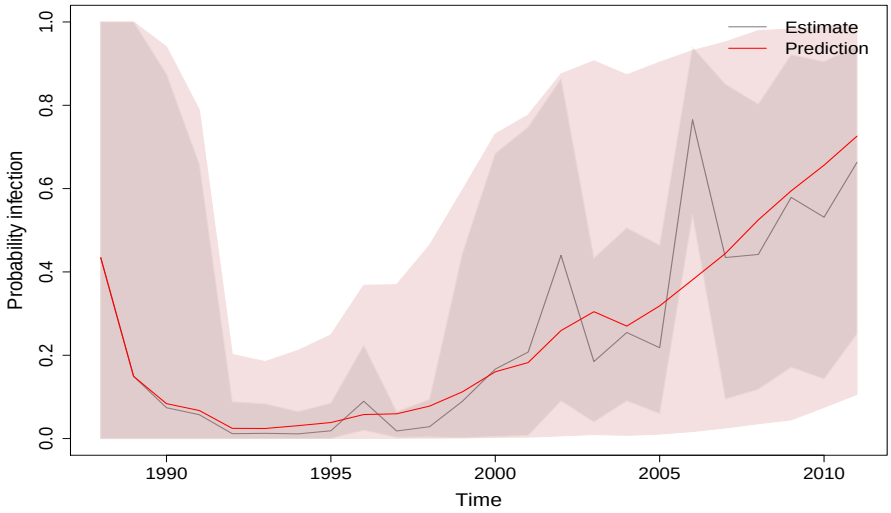

Figure S3: Comparison between the mean and 95% credible intervals of the estimated (grey) and predicted dog probability of infection (red) from the CPV dynamics model

S3. JAGS code

The JAGS code for the two models can be found in a separate supplementary file associated with this submission.

Note: *Ntime* (total number of years in time-series), *Ndogs* and *Nlions* (total of number of dogs and lions sampled respectively), *V_regional* (regional-level vaccination coverage; i.e. proportion of vaccinated villages per year), *Coverage* (village-level vaccination coverage), *Birth_sp* (year of birth of each individual *i*, *sp* corresponding to *dogs* or *lions*), *Sample_sp* (year when the serum sample of individual *i* was collected), and *titer* (titre binary result where 1 indicates positive for CDV infection and 0 indicates negative) correspond to available data.

S4. Results

As indicated by the normal posteriors distributions, the models investigated converged well. In addition, Tables S5-6 show the parameter values resulting from the models and the estimated shrink factor values of the Gelman-Rubin convergence diagnostic. If both the estimated value and the upper credible interval of the shrink factor are close to 1, the chains are considered to have converged. As a rule of thumb, a model is considered to be appropriate if approximately 80% of the parameter estimates converged well [16].

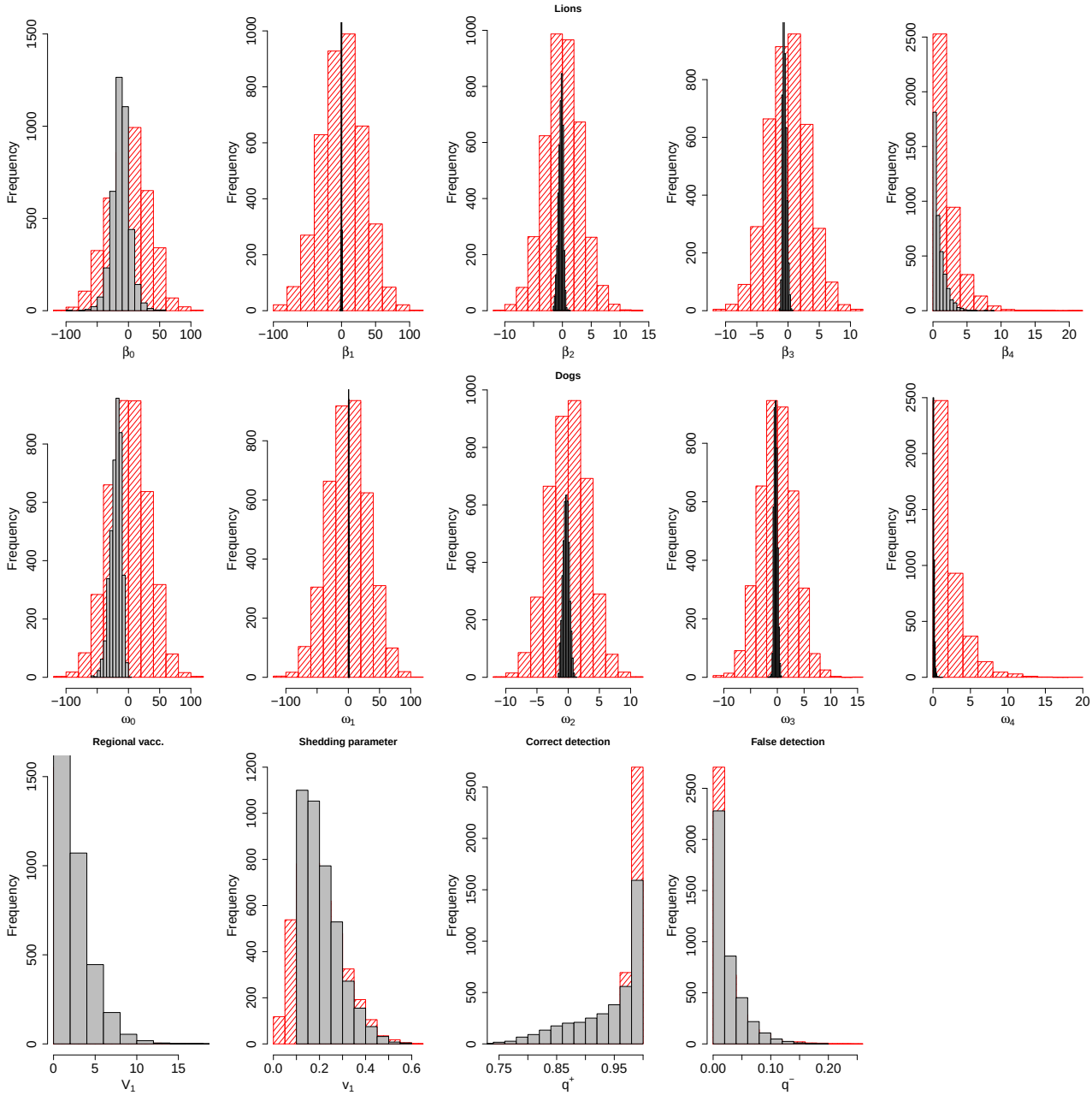

Figure S4: Histograms of the prior distributions (red) and posterior distributions (grey) of the domestic-wildlife interface model.

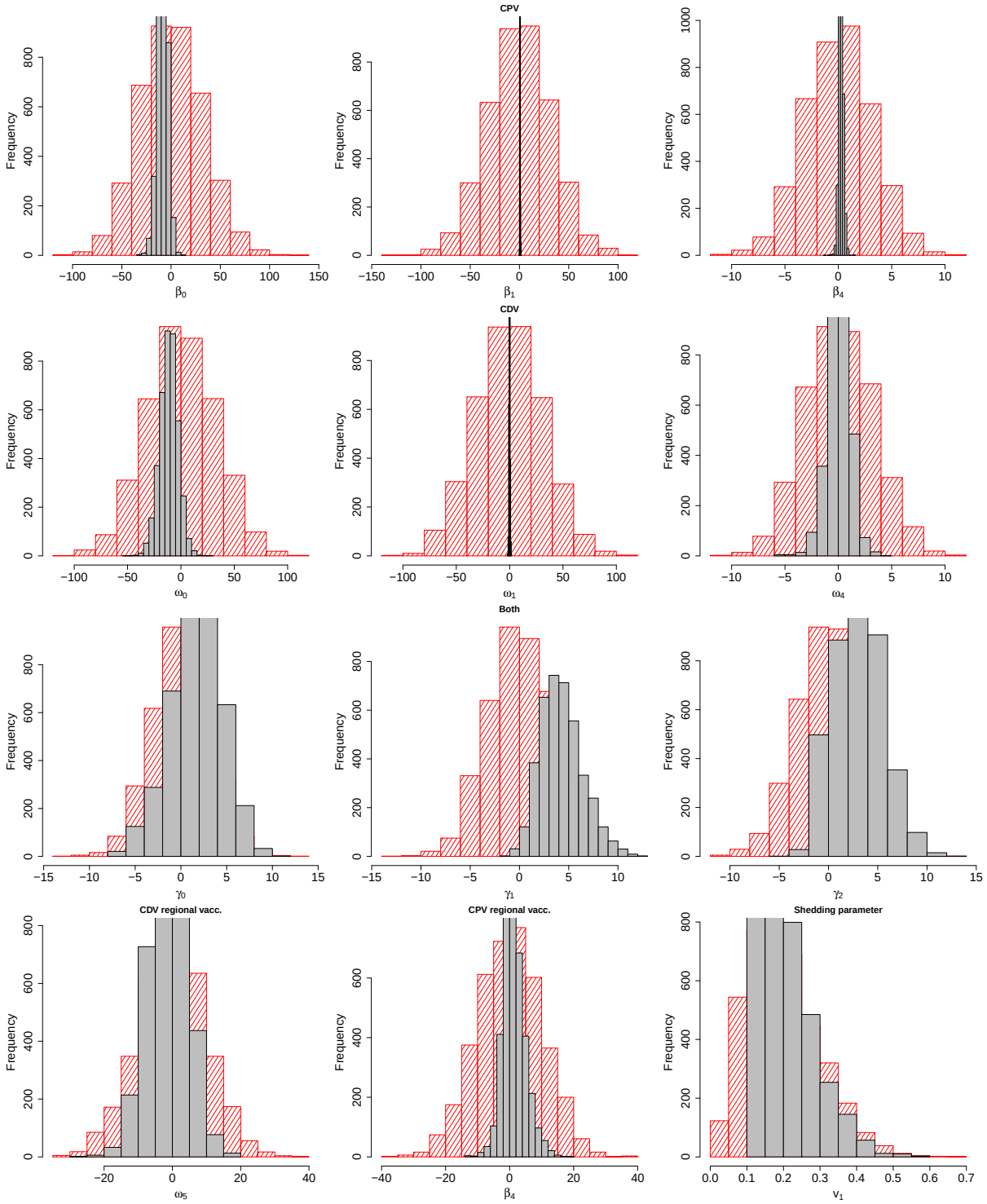

Figure S5: Histograms of the prior distributions (red) and posterior distributions (grey) of the co-infection in dogs model.

Table S5: Median posterior estimates with associated 95% credible intervals (CI) and Gelman-Rubin shrink factor (estimated value [upper CI]) of the domestic-wildlife interface model.

Species	**Variable**	**Param.**	**Median (95% CI)**	**Shrink factor**
Lions	Intercept	β_0	-11.818 (-40.65, 15.19)	1.02 [1.10]
	Linear trend	β_1	-0.039 (-0.840, 0.788)	1.02 [1.12]
	AR(1)	β_2	-0.246 (-1.125, 0.381)	1.00 [1.01]
	AR(2)	β_3	-0.567 (-1.013, 0.129)	1.01 [1.02]
	Dog-to-lion transmission	β_4	0.580 (0.019, 3.277)	1.02 [1.09]
	Correct detection	q^+	0.966 (0.798, 1.00)	1.01 [1.03]
	False detection	q^-	$1.545{\cdot}10^{-2}$ ($2.77{\cdot}10^{-5}$, $9.71{\cdot}10^{-2}$)	1.00 [1.01]
Dogs	Intercept	ω_0	-19.02 (-39.33, -6.09)	1.06 [1.18]
	Linear trend	ω_1	0.557 (0.188, 1.153)	1.07 [1.19]
	AR(1)	ω_2	-0.334 (-1.24, 0.580)	1.06 [1.24]
	AR(2)	ω_3	-0.312 (-0.799, 0.311)	1.00 [1.01]
	Lion-to-dog transmission	ω_4	0.072 (0.003, 0.324)	1.00 [1.02]
	Regional vacc. (lag 1)	ω_5	1.71 (0.071, 7.63)	1.00 [1.00]
	Shedding parameter	v_1	0.193 (0.105, 0.407)	1.00 [1.00]
	Correct detection	q^+	0.914 (0.87, 0.95)	1.00 [1.00]
	False detection	q^-	$4.19{\cdot}10^{-3}$ ($1.09{\cdot}10^{-5}$, $4.00{\cdot}10^{-2}$)	1.01 [1.03]

Table S6: Median posterior estimates with associated 95% credible intervals (CI) and Gelman-Rubin shrink factor (estimated value [upper CI]) of the co-infection model.

Disease	**Variable**	**Param.**	**Median (95% CI)**	**Shrink factor**
CPV	Intercept	β_0	-7.974 (-19.44, -1.17)	1.11 [1.39]
	Linear trend	β_1	0.570 (0.104, 1.180)	1.03 [1.24]
	AR(1)	β_2	-0.162 (-0.840, 0.549)	1.04 [1.10]
	AR(2)	β_3	-0.171 (-0.721, 0.315)	1.10 [1.38]
	CDV-to-CPV transmission	β_4	0.255 (-0.137, 0.687)	1.70 [3.00]
	Regional vacc. (lag 1)	β_5	0.98 (-4.83, 9.50)	1.09 [1.35]
	Shedding parameter	v_1	0.188 (0.104, 0.39)	1.00 [1.00]
	Correct detection	q^+	0.973 (0.917, 1.00)	1.05 [1.22]
	False detection	q^-	$3.11{\cdot}10^{-3}$ ($8.56{\cdot}10^{-6}$, $5.61{\cdot}10^{-2}$)	1.05 [1.16]
CDV	Intercept	ω_0	-10.97 (-28.63, 4.82)	1.11 [1.37]
	Linear trend	ω_1	-0.114 (-0.97, 0.91)	1.12 [1.42]
	AR(1)	ω_2	-0.348 (-1.03, 0.27)	1.02 [1.05]
	AR(2)	ω_3	-0.063 (-0.74, 0.707)	1.06 [1.22]
	CPV-to-CDV transmission	ω_4	0.201 (-2.08, 1.93)	1.09 [1.32]
	Regional vacc. (lag 1)	ω_5	-1.28 (-12.66, 9.66)	1.02 [1.07]
	Correct detection	q^+	0.960 (0.85, 1.00)	1.47 [3.08]
	False detection	q^-	$8.73{\cdot}10^{-4}$ ($2.54{\cdot}10^{-6}$, $5.22{\cdot}10^{-3}$)	1.15 [1.54]
Both	Independent infections	γ_0	-1.787 (-4.504, 7.05)	2.17 [4.33]
	CPV-to-coinfection	γ_1	4.098 (0.673, 9.041)	2.17 [4.40]
	CDV-to-coinfection	γ_2	2.98 (-1.30, 8.17)	1.04 [1.15]
	Correct detection	q^+	0.989 (0.92, 1.00)	1.10 [1.30]
	False detection	q^-	$2.691{\cdot}10^{-4}$ ($8.14{\cdot}10^{-7}$, $2.65{\cdot}10^{-3}$)	1.04 [1.07]

Comparison of the estimated temporal profiles and the observed data

Fig. S6 shows the estimated CPV temporal profiles for the dogs and the lions, overlaid on the observed data. Despite the large confidence intervals as well as the small sample size (in particular for the lions, as seen with the standard error bars in fig. S6), our model is able to properly determine the temporal position of the peaks of infection.

Our results also show that our model is capable to capture the general trend of CPV infection for both the lions and the dogs, inferring the temporal profile for periods of time when no data is available (*cf* dogs between 1998 and 2002).

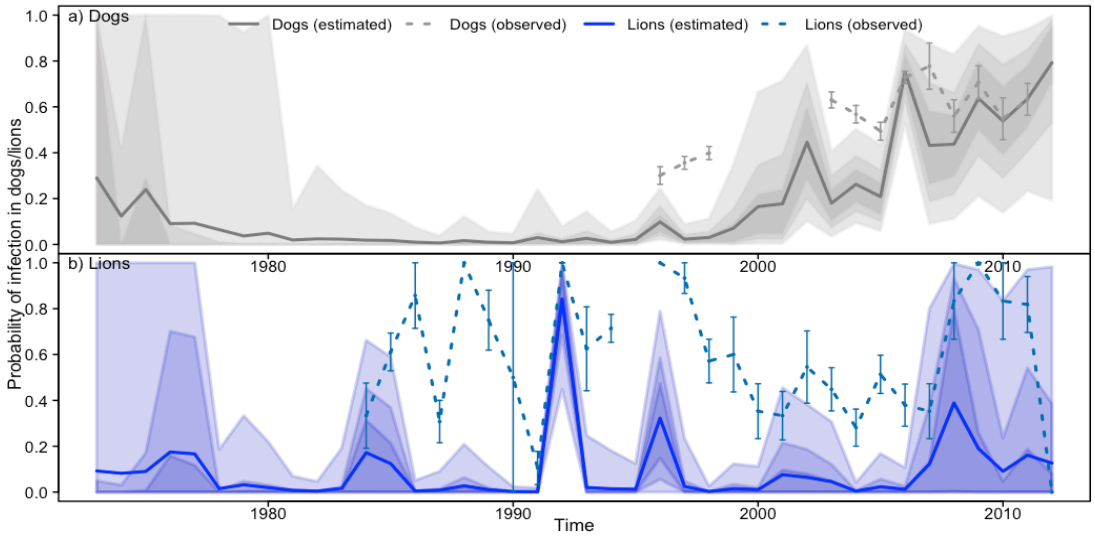

Figure S6: Estimated and observed CPV temporal profiles. a) Mean estimated annual probability of CPV infection (solid line) and observed prevalence (dotted line) in dogs. b) Mean estimated annual probability of CPV infection (solid line) and observed prevalence (dotted line) in lions. Associated 50%, 75% and 95% credible intervals shown in increasingly lighter colour shading. The vertical lines correspond to the standard error for the observed prevalence.

Age-seroprevalence profiles

Age-seroprevalence temporal profiles were generated for lions and dogs from the data to clarify the potential peaks of infection. In Fig. S7 and Fig. S8 each curve corresponds to a different sampling period, and represents seroprevalence in the birth year, as opposed to the sampling year. This attempts to capture the fact that animals born after a peak of infection will likely only get infected during the next peak. Thus for a given sampling period, if the youngest age-class presents a low seroprevalence, it suggests that they were born after an outbreak. If in a directly following sampling period, the lions born in similar years present a high prevalence, it suggests the outbreak occurred between those two sampling periods.

For example, we show that for the 1984-1991 sampling period, lions born between 1985 and 1990 present a very low sero-prevalence. However during the following sampling period (1992-1995), the lions born in the same years (1985 to 1990) present a high sero-prevalence. This clearly suggests that an outbreak occurred around 1991. Following the same reasoning, CPV outbreaks in lions seem to have occurred around 1995 and around 2001.

In the case of endemic infection, we wouldn't observe any peak, since the infection would be constant in time. In that scenario, the age-seroprevalence curves would be constant, or at least show limited variations corresponding to the absence of marked peaks of infection. On the opposite, fig. S8 presents systematic decreases in sero-prevalence at the end of each sampling period, suggesting the presence of peaks of infection followed by periods of low infection. These results thus support the hypothesis of non-endemic infection in lions.

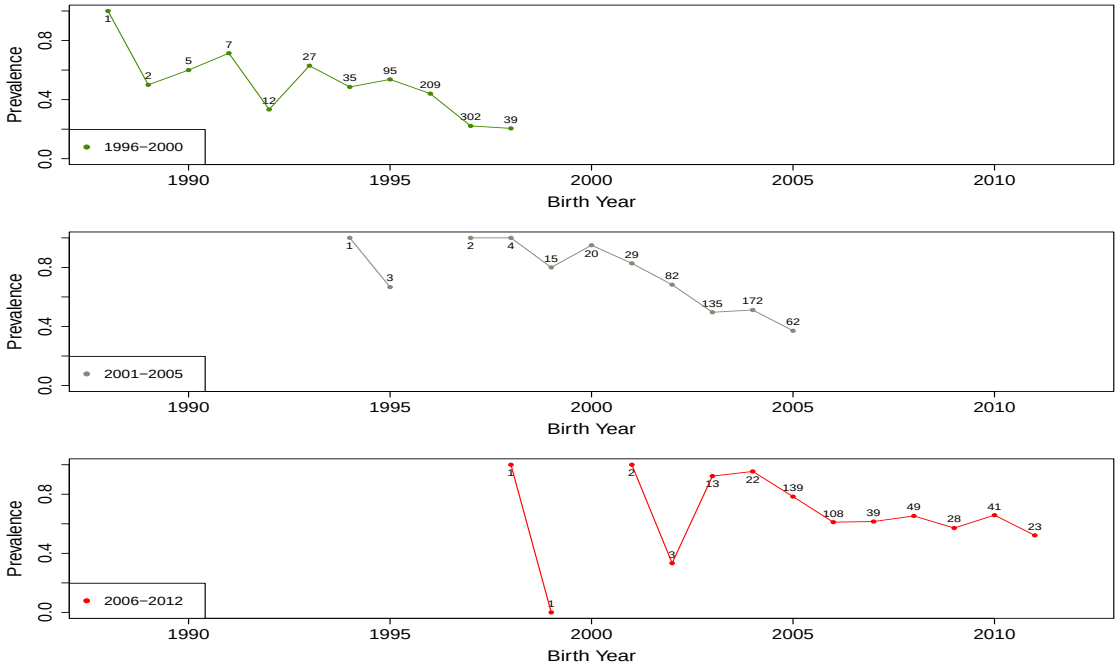

Figure S7: Dog age-seroprevalence temporal profiles. From top to bottom, panels correspond to data sampled between 1984-1991, 1992-1995, 1996-2001, 2002-2007 and 2008-2012. The number indicate the sample sizes for each point.

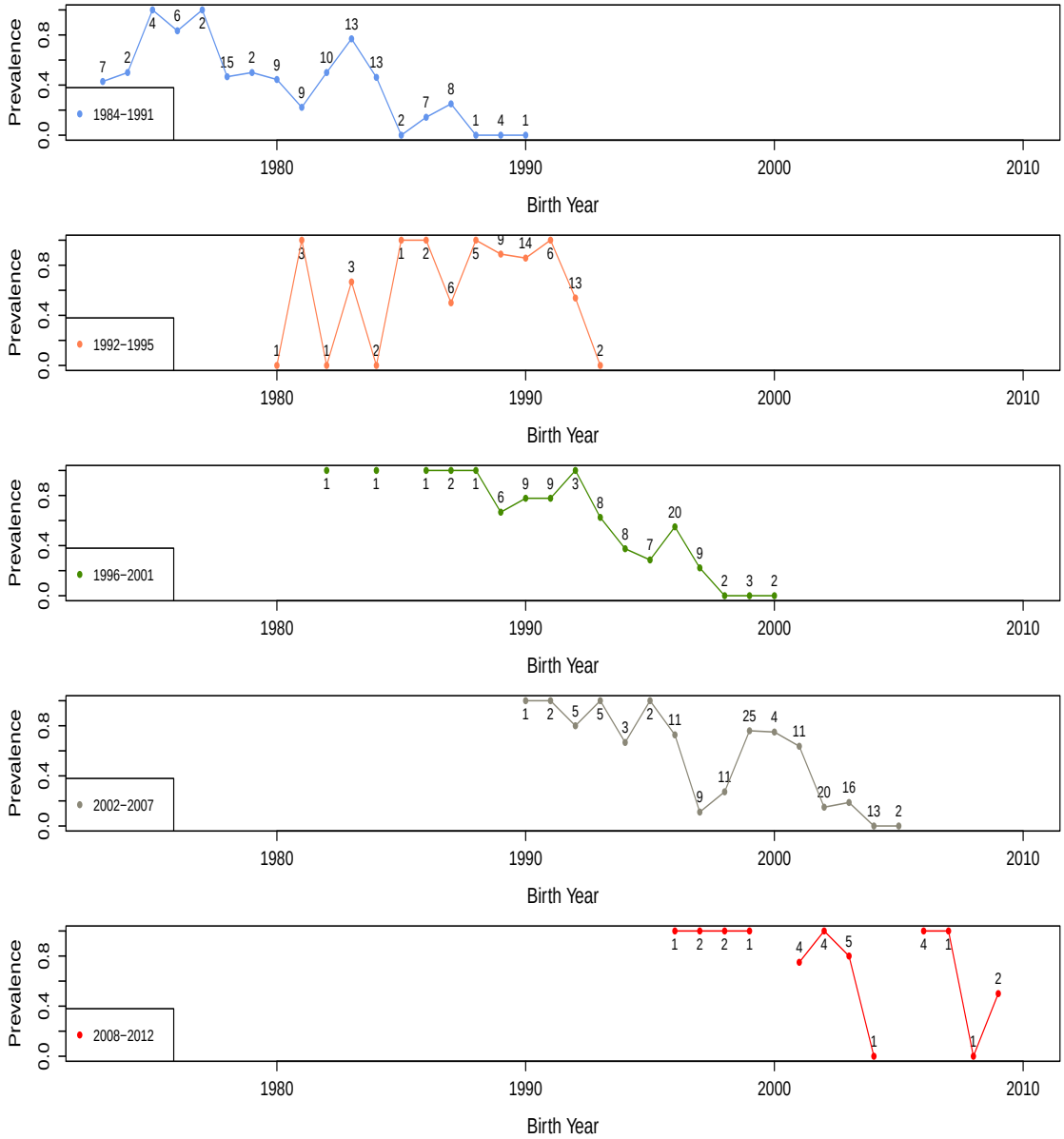

Figure S8: Lion age-seroprevalence temporal profiles. From top to bottom, panels correspond to data sampled between 1984-1991, 1992-1995, 1996-2001, 2002-2007 and 2008-2012. The number indicate the sample sizes for each point.

Cyclicity of CPV in lions and dogs

A Royama triangle is used to identify cyclicity in a time-series generated from an autoregressive process (Figure S9). If the estimated AR1 and AR2 coefficients fall outside the triangle when plotted against each other, it indicates unstable dynamics that become extinct. Inside the triangle, the dynamics are stable or display damped oscillations. In addition, if the parameters fall within the parabola, the dynamics are deemed cyclic with the period increasing from left to right as represented by the contour lines. Our results (Figure S9) shows that CPV dynamics are cyclic in both species, with a period of ∼3 to 4 years between peaks of infection.

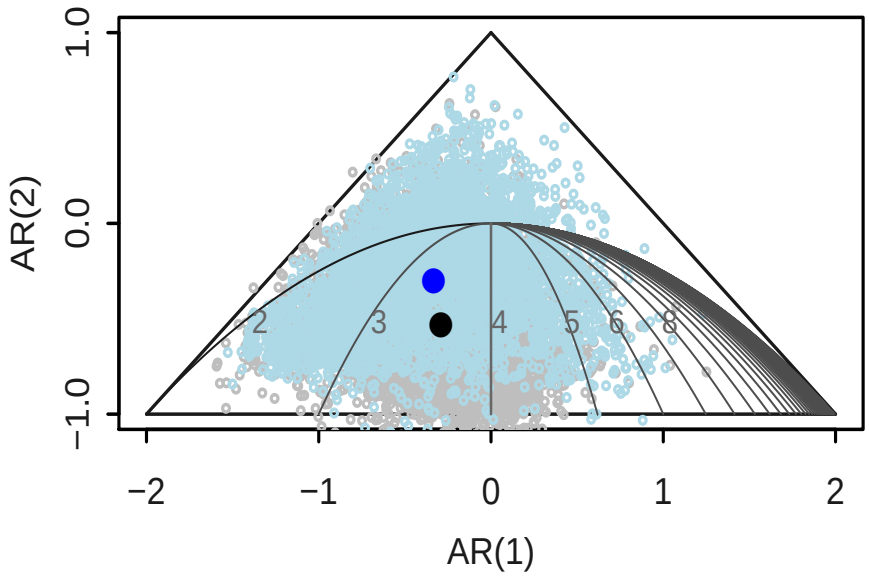

Figure S9: AR parameter plane. Closed dark dots correspond to the posterior median of the first-order AR coefficient [x-axis, AR (1)] of model 2 (i.e.,ω_2/β_2) against the posterior median of the second-order AR coefficient [y-axis, AR (2)] (i.e., ω_2/β_2) for dogs (black) and lions (blue), respectively. Each open light dot corresponds to a draw of the AR posterior distributions.

Dynamics of CPV in Dogs without shedding

Figure S10a shows that without including shedding, the model for CPV infection in domestic dogs predicts higher probabilities of infection (since it combined the natural and vaccine shed infection) but with similar dynamics. Moreover, vaccination can decrease the annual proability of CPV infection, as shown in S10b. In particular, we note an increase of the difference with/without vaccination after 2003, when the vaccination effort is more important.

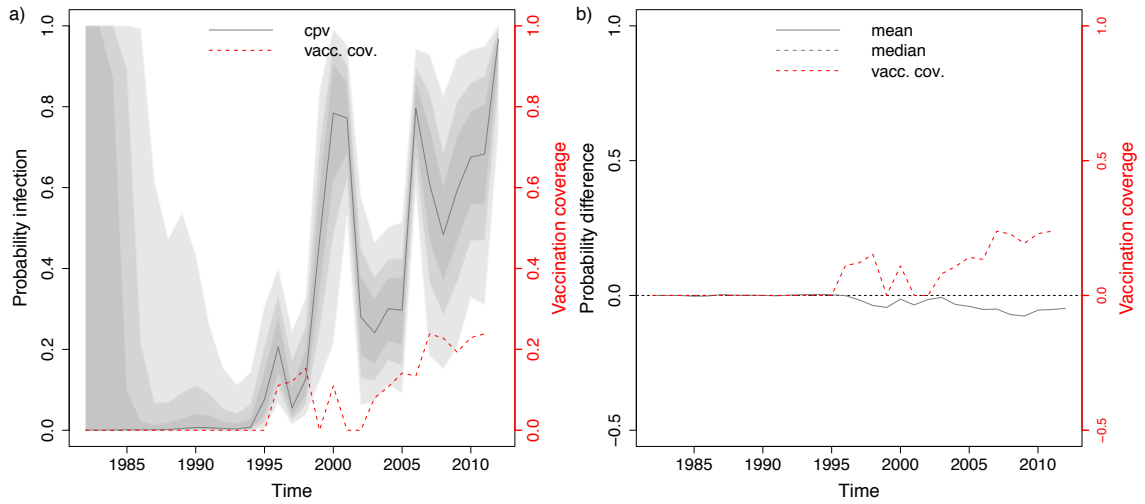

Figure S10: a) Median annual probability of CPV infection in domestic dogs (full line) and vaccination coverage (red). b) Mean (full line) and median (grey dotted line) difference between the estimated prediction of the annual probability of CPV infection and the predicted annual probability of CPV infection without the vaccination effect. Associated 50%, 75% and 95% credible intervals shown in increasingly lighter colour shading. Vaccination coverage shown in red.

Influence of CPV on co-infection dynamics and effect of vaccination.

The influence of both CDV and CPV on the co-infection dynamics starts decreasing in 2003 (Fig. 5, main text). This suggests that the vaccination effort, which increases in 2003, might have decoupled both diseases. This result is further confirmed when we compare the estimated probability of being infected by CPV and CDV (p_{both}) to the theoretical probability in the case where both diseases are independent, which is simply the product of both probabilities $p_{CDV} * p_{CPV}$ (fig. S11). This difference is negative until 2005 when it starts increasing, which means that before 2005, more dogs are positive to both diseases than expected when we consider both the diseases separately. However after 2005 (fig.S11, vertical dashed line), this trend is reversed: the increase in the vaccination efforts tends to decouple the two diseases.

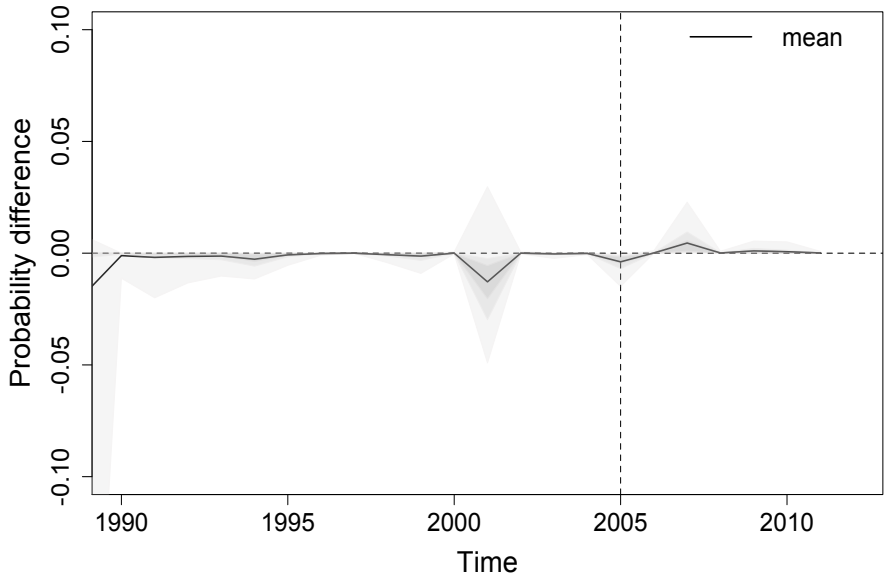

Figure S11: Mean difference between the estimated probability of being infected by both diseases and the probability of infection under strict independence of CPV and CDV (grey line). Credible intervals (95, 75 and 50%) shown in increasing lighter shades of grey.

References

0. Viana M, et al. (2015) Dynamics of a morbillivirus at the domestic-wildlife interface: Canine distemper virus in domestic dogs and lions. Proceedings of the National Academy of Sciences 112.5:1464-1469.
1. Pennycuick CJ & Rudnai J (1970) A method of identifying individual lions *Panthera leo* with an analysis of the reliability of identification. *J Zool* 160(4):497-508.
2. Packer C & Pusey AE (1993) Should a lion change its spots? *Nature* 362(6421):595.
3. Pusey AE & Packer C (1994) Non-offspring nursing in social carnivores: minimizing the costs. *Behav Ecol* 5(4):362-374.
4. Whitman KL, Starfield AM, Quadling H, & Packer C (2007) Modeling the effects of trophy selection and environmental disturbance on a simulated population of African lions. *Conserv Biol* 21(3):591-601.
5. Packer C, et al. (1999) Viruses of the Serengeti: patterns of infection and mortality in African lions. *J Anim Ecol* 68(6):1161-1178.
6. Kaare M, et al. (2009) Rabies control in rural Africa: Evaluating strategies for effective domestic dog vaccination. Vaccine 27(1):152-160.
7. Cleaveland S (1996) The epidemiology of rabies and canine distemper in the Serengeti, Tanzania. PhD Thesis (University of London).
8. Halliday JEB (2010) Animal sentinel surveillance: Evaluating domestic dogs as sentinels for zoonotic pathogen surveillance. PhD Thesis. University of Edinburgh.
9. Almberg ES, Mech LD, Smith DW, Sheldon JW, & Crabtree RL (2009) A serological survey of infectious disease in Yellowstone National Parks canid community. *PLoS ONE* 4(9):e7042.
10. Gowtage-Sequeira S, et al. (2009) Epidemiology, pathology, and genetic analysis of a canine distemper epidemic in Namibia. *J Wildl Dis* 45(4):1008-1020.
11. Prager KC, et al. (2012) Rabies virus and canine distemper virus in wild and domestic carnivores in northern Kenya: are domestic dogs the reservoir? *EcoHealth* 9(4):483-498.
12. Carmichael LE, Joubert JC, & Pollock RV (1980) Hemagglutination by canine parvovirus: serologic studies and diagnostic applications. *American J Vet Res* 41(5):784-791.

13. Gilbert AT, et al. (2013) Deciphering serology to understand the ecology of infectious diseases in wildlife. *EcoHealth* 10(3):298-313.
14. Gelman A & Rubin D (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* 7:457-511.
15. Brooks SP & Gelman A (1998) General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 7:434-455.
16. Spiegelhalter DJ, Best NG, Carlin BP, & van der Linde A (2002) Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 64(4):583-639. 21.