

The following supplement accompanies the article

Phylogenomic insights to the origin and spread of phocine distemper virus in European harbour seals in 1988 and 2002

Iben Stokholm*, Tero Härkönen, Karin Harding, Ursula Siebert, Kristina Lehnert, Rune Dietz, Jonas Teilmann, Anders Galatius, Linnea Worsøe Havmøller, Emma Carroll, Ailsa Hall, Morten Tange Olsen*

*Corresponding authors: morten.olsen@snm.ku.dk

Diseases of Aquatic Organisms 133: 47–56 (2018)

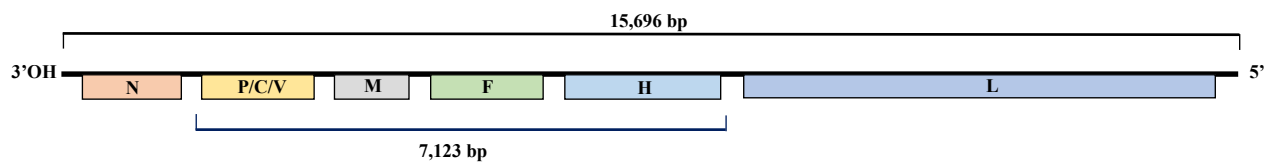


Fig. S1. A schematic drawing of the genome of Phocine Distemper Virus (PDV) including the six non-overlapping genes; N (position 108-1,679), P (position 1,801-3,324), M (position 3,432-4,439), F (position 5,028-6,923), H (position 7,079-8,902) and L (position 9,036-15,590) coding for at least eight proteins as depicted (de Vries *et al.*, 2013) sequenced in this study (position 1,744-8,915).

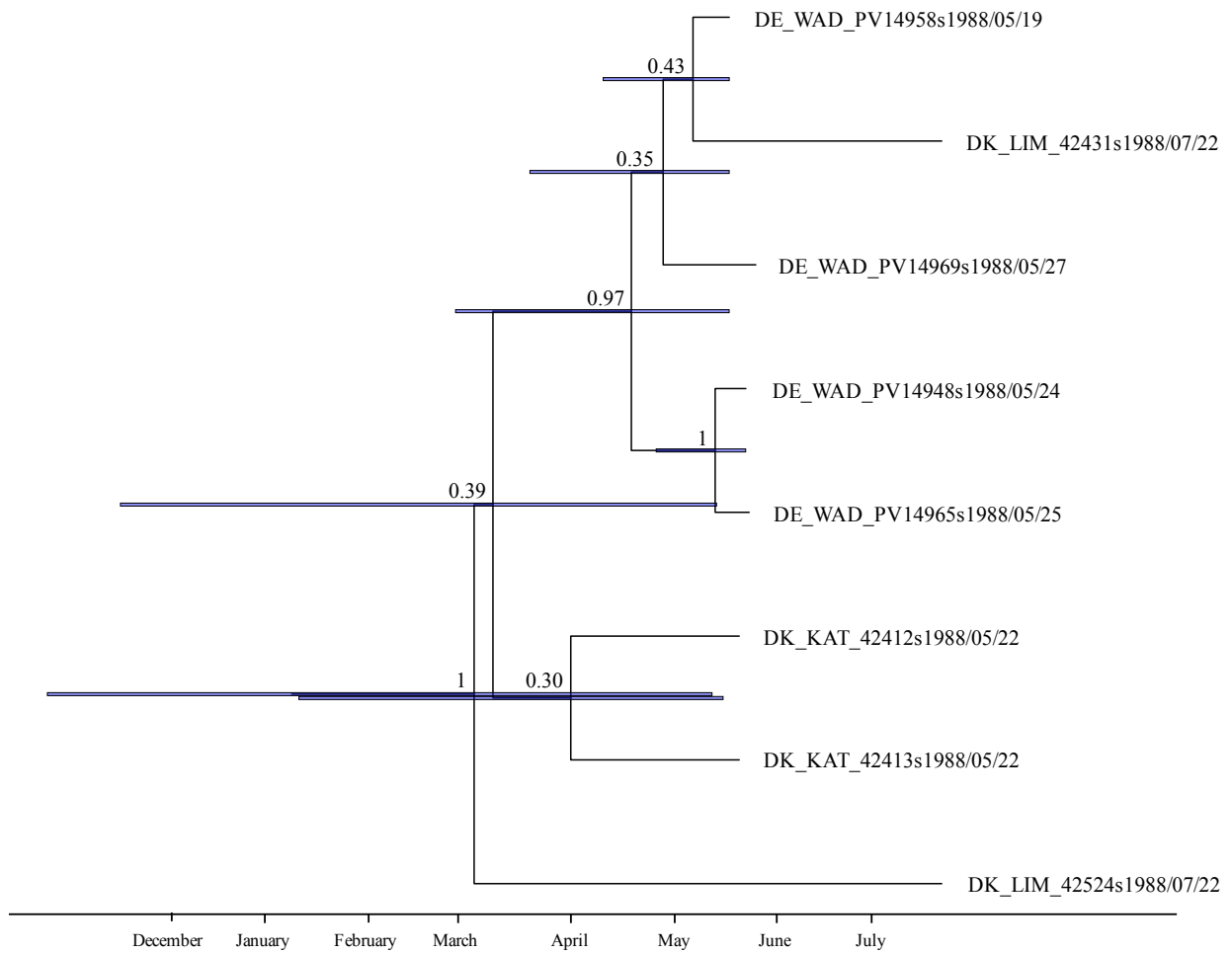


Fig. S2. Phylogenetic relationship among eight PDV strains sequenced from harbour seals infected during the 1988 epidemic. The analysis suggests at least two independent introductions of the PDV strain to Limfjorden (DK_LIM_42524; 1988/07/22) and Kattegat (DK_KAT_42412; 1988/05/22), respectively. Node bars are shown as the 95 % interval of the node heights thereby indicating their uncertainty.

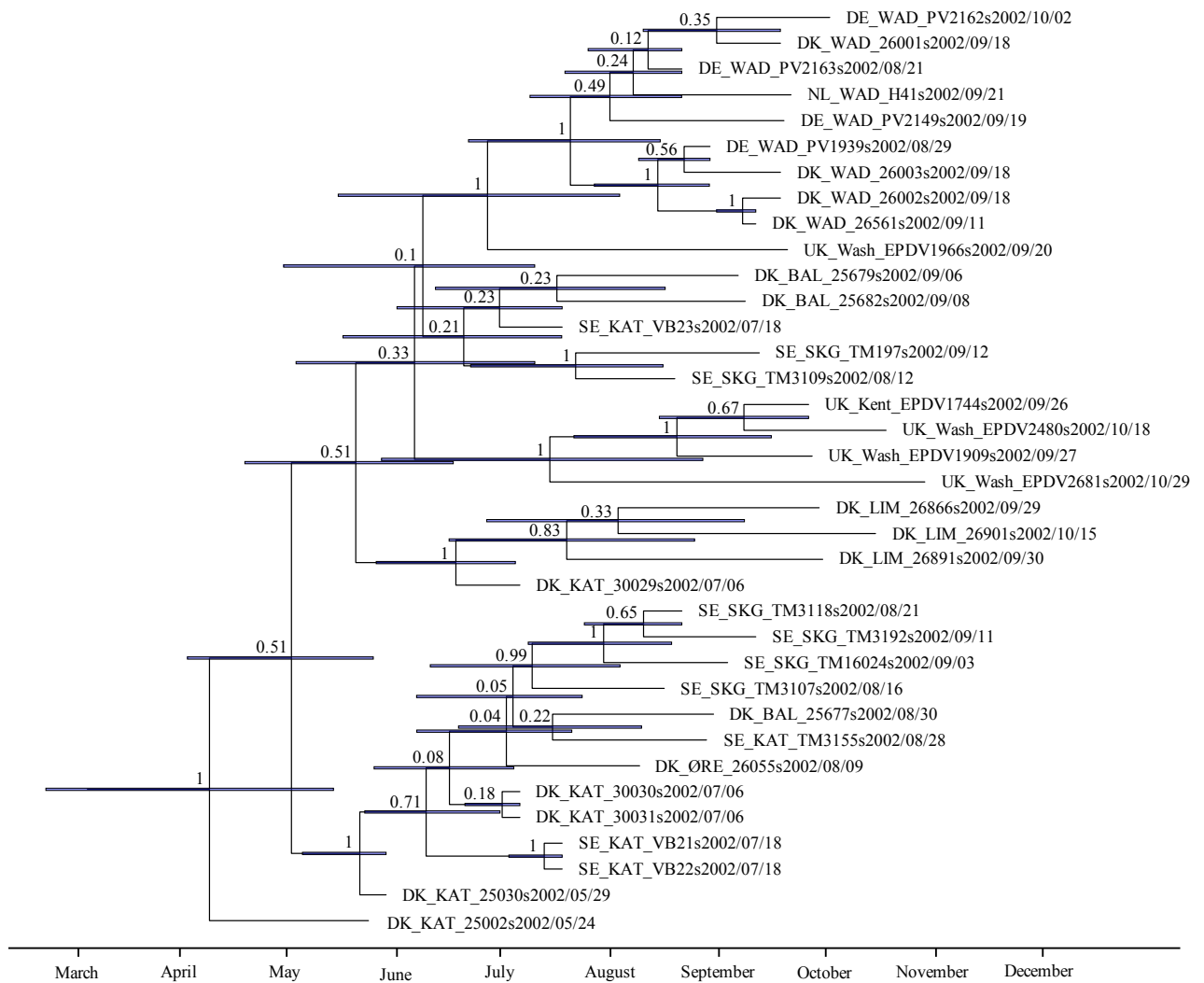


Fig. S3. Phylogenetic relationship among PDV strains sequenced from harbour seals infected during the 2002 epidemic. The analysis provides strong support for Kattegat as the origin of the outbreak (DK_KAT_25002; 2002/05/24), with subsequent pulses of spread to harbour seal localities in Skagerrak (SKG), SW Baltic (BAL), the Danish German and Dutch Wadden Sea (WAD), Limfjorden (LIM) and the Wash (UK_Wash). Node bars are shown as the 95 % interval of the node heights thereby indicating their uncertainty.

Table S1. Summary of tissue type, storage conditions and the number of samples extracted and included in the analyses.

Country	Tissue type		Storage temperature		Sample storage		Number of samples		Number of samples included in the data set	
Denmark	Liver	Spleen/lung	-20 °C	-20 °C	Organs in bags	Organs in bags	27	28	4	16
England	-	Spleen/lung	-	-20 °C/-80 °C	-	-	-	17		5
Germany	Muscle	Spleen/lung	-20 °C	-20 °C	Organs in bags	Organs in bags	5	5	4	4
The Netherlands	-	Muscle	-	-20 °C	-	-	-	2		1
Scotland	-	Spleen/lung	-	-20 °C	-	-	-	4		0
Sweden	-	Spleen	-	-20 °C/-80 °C	-	Tubes RNAlater	-	11		10
Total							32	67	8	36

Table S2. Table of primers used to amplify the fragments sequenced. The primers were compiled from several previous studies and combined to construct the complete genome sequence of PDV (de Vries et al., 2013).

Primer	Sequence	Position start	Position end	Coverage base pairs	Gene
4	5'-TCAGCTAAGTTCTAAAAGACATTGC-3' 3'-CAGTTAGATGAAGCATTTCTTCGG-5'	1.672	2.303	631	P
5	5'-GAAGGAAGAGAGGGAAGCCTTGATG-3' 3'-GTTTTTTGATCGAGTCAATCTCTCC-5'	2.224	2.854	630	P
6	5'-TGAGGATAATCAATCAATACTTAGC-3' 3'-TGAGTCCTAAGTTTTTTGTAATTGC-5'	2.778	3.407	629	P
7	5'-ACCAATGCCTTTGTTAGTAATCAGG-3' 3'-TCGGAAATCAAACATTCGCGAGGG-5'	3.333	3.974	641	P/M
8	5'-GTAGTGTACATGAGTATTACTAGAC-3' 3'-AACAGACTAAAGCTAATTAATGCTG-5'	3.9	4.549	649	M
9	5'-TAATGATATAGACTGAAGCACTCAC-3' 3'-TTCAGGGTTTAGAAAATGGTGAATCG-5'	4.464	5.108	644	M/F
10	5'-ATGACTCGAGTCAAGAACTCCCAG-3' 3'-CAGCAGTTGCTACACCTAGAGCTGC-5'	5.028	5.655	627	F
11	5'-GGGATCAGGTAGAAGACAGAGGCG-3' 3'-GGGACAGTAGTGTACCATTCCCTGTG-5'	5.582	6.217	635	F
12	5'-GGGTTGTCGTACATAGATTGGAGGC-3' 3'-ATACCGGGAGTGCAATGAGGGTGCC-5'	6.148	6.777	629	F
13	5'-AATCCTCTGATCAAATCCTTGACAC-3' 3'-ATCCTTGACTTGATGATGCACTGCC-5'	6.703	7.348	645	F/H
14	5'-AAAAAGCAATTTGGAGTTCAATAAG-3' 3'-GATAATCCTGTAGTTTGTGTTTGG-5'	7.27	7.9	630	H
H5/H6	5'-AGATGATATCTTTCCTCC-3' 3'-ATCCATATGAGTTGCTCC-5'	7.615	8.065	451	H
15	5'-GATTCGTGTCTTTGAGATTGGATTG-3' 3'-CCATTTTTAGGTGGAATTGTTAGCC-5'	7.819	8.448	629	H
16	5'-ATAATATTTAACGGTGATGGGATGG-3' 3'-AAACAAGGGAAAGAAACACTTACCG-5'	8.366	8.972	606	H

Table S3. RT-PCR conditions for the 14 different primer sets used in this study.

Cycles	1	1	40 / 38	x (1	1	1)	1	1
Time (min)	30	15		1	1	1	10	∞
Primer 4 (°C)	50	95	x	94	64	72	72	4
Primer 5 (°C)	50	95	x	94	64	72	72	4
Primer 6 (°C)	50	95	x	94	63	72	72	4
Primer 7 (°C)	50	95		x	94	65	72	4
Primer 8 (°C)	50	95	x	94	63	72	72	4
Primer 9 (°C)	50	96	x	94	62	72	72	4
Primer 10 (°C)	50	95	x	94	62	72	72	4
Primer 11 (°C)	50	95	x	94	62	72	72	4
Primer 12 (°C)	50	95	x	94	62	72	72	4
Primer 13 (°C)	50	95	x	94	62	72	72	4
Primer 14 (°C)	50	95	x	94	58	72	72	4
Primer H5/H6 (°C)	50	95	x	94	55	72	72	4
Primer 15 (°C)	50	95	x	94	58	72	72	4
Primer 16 (°C)	50	95		x	94	64	72	4

Table S4. The final sample set of PDV sequences covering the coding and non-coding regions of the P, M, F and H genes of samples collected in 2002 and 1988. Numbers 1-4 indicate identical sequences. * Note that viral RNA was not successfully sequenced with primer 6 for the samples 42412 and 42413 and a fragment from a German sample was used to fill the gap (see Table 9).

Sample	Species	Location	Area	Date	Tissue
25002	Harbour seal	Læsø	DK-KAT	24. May 2002	(Liver)
25030 ⁴	Harbour seal	Anholt	DK-KAT	29. May 2002	Spleen
30029	Harbour seal	Samsø	DK-KAT	6. July 2002	Spleen
30030 ⁴	Harbour seal	Samsø	DK-KAT	6. July 2002	Spleen
30031 ⁴	Harbour seal	Samsø	DK-KAT	6. July 2002	Spleen
VB21 ³	Harbour seal	Prästaskär	SE-KAT	18. July 2002	Spleen
VB22 ³	Harbour seal	Yttre Flatbåden	SE-KAT	18. July 2002	Spleen
VB23	Harbour seal	Prästaskär	SE-KAT	18. July 2002	Spleen
TM3155	Harbour seal	Luseskär	SE-KAT	28. August 2002	Spleen
TM197	Harbour seal	Väderöarna	SE-SKG	12. September 2002	Spleen
TM3107	Harbour seal	Sälvik	SE-SKG	16. August 2002	Spleen
TM3109	Harbour seal	Timglaset, Koster	SE-SKG	19. August 2002	Spleen
TM3118	Harbour seal	Lilla Sockna	SE-SKG	21. August 2002	Spleen
TM3192	Harbour seal	Tjärnö	SE-SKG	11. September 2002	Spleen
TM16024	Harbour seal	Ramsökälven	SE-SKG	3. September 2002	Spleen
26055	Harbour seal	Karlslunde	DK-ØRE	9. August 2002	Spleen
25679	Harbour seal	Møn	DK-BAL	6. September 2002	Spleen
25682	Harbour seal	Rødsand	DK-BAL	8. September 2002	Spleen
25677	Harbour seal	Vester Kippinge	DK-BAL	30. August 2002	Spleen
H41	Harbour seal	Ameland	NL-WAD	9. September 2002	Muscle
Pv 1939	Harbour seal	Eiderstedt	DE-WAD	29. August 2002	Spleen/lung
Pv 2149	Harbour seal	Eidermündung	DE-WAD	19. September 2002	Spleen/lung
Pv 2162	Harbour seal	Sylt	DE-WAD	2. October 2002	Spleen
Pv 2163 ²	Harbour seal	Eiderstedt	DE-WAD	21. August 2002	Spleen
26001 ²	Harbour seal	Rømø (North)	DK-WAD	18. September 2002	Spleen
26002 ¹	Harbour seal	Rømø (North)	DK-WAD	18. September 2002	Spleen
26003	Harbour seal	Rømø (North)	DK-WAD	18. September 2002	Spleen
26561 ¹	Harbour seal	Lindet, Koresand	DK-WAD	11. September 2002	Spleen
26866	Harbour seal	Blinde Røn	DK-LIM	29. September 2002	Spleen
26891	Harbour seal	Blinde Røn	DK-LIM	30. September 2002	Spleen
26901	Harbour seal	Blinde Røn	DK-LIM	15. October 2002	Spleen
EPDV2681	Harbour seal	Wash	UK-Wash	29. September 2002	Spleen/lung
EPDV2480	Harbour seal	Wash	UK-Wash	18. October 2002	Spleen/lung
EPDV1966	Harbour seal	Wash	UK-Wash	20. September 2002	Spleen/lung
EPDV1909	Harbour seal	Wash	UK-Wash	27. September 2002	Spleen/lung
EPDV1744	Harbour seal	Kent	UK-Wash	26. September 2002	Spleen/lung
PV14965	Harbour seal	Wadden Sea	DE-WAD	25. May 1988	Muscle/blubber
PV14958	Harbour seal	Wadden Sea	DE-WAD	19. May 1988	Muscle/blubber
PV14948	Harbour seal	Wadden Sea	DE-WAD	24. May 1988	Muscle/blubber
PV14969	Harbour seal	Wadden Sea	DE-WAD	27. May 1988	Muscle/blubber
42524	Harbour seal	Limfjorden	DK-LIM	22. July 1988	Liver
42531	Harbour seal	Limfjorden	DK-LIM	22. July 1988	Liver
42412 *	Harbour seal	Anholt	DK-KAT	22. May 1988	Liver
42413 *	Harbour seal	Anholt	DK-KAT	22. May 1988	Liver

Table S5. Results of the substitution saturation tests of the alignment of PDV sequences from 1988 and 2002. No substitution saturation is detected if the Iss is significantly lower than Iss.c (Xia et al. 2003).

Codon(s)	Iss	Iss.c
Codon position 1 and 2	0.0117	0.8026
Codon position 3	0.0027	0.7909

Table S6. Results of the substitution model tests conducted in JModeltest (Posada 2008). The score is the weights of AIC, AICc, BIC and DT and the model of the highest weight is the model which is best fitted for the given alignment.

Alignment	AIC		AICc		BIC		DT	
	Model	Score	Model	Score	Model	Score	Model	Score
PDVall	TVM + I	0.164	TPM1uf + I	0.165	HKY + I	0.307	HKY + I	0.016
PDV1988	TPM1uf	0.302	TPM1uf	0.304	HKY	0.487	HKY	0.057
PDV2002	TPM1uf + I	0.137	TPM1uf + I	0.138	HKY	0.731	HKY	0.031

Table S7. General priors applied to the parameters in BEAUTi.

Parameter	Distribution	Initial value	Interval
gammaShape.s:1stpos	Exponential	1	$[-\infty, \infty]$
gammasShape.s:2ndpos	Exponential	1	$[-\infty, \infty]$
gammasShape.s:3rdpos	Exponential	1	$[-\infty, \infty]$
gammasShape.s:coding	Exponential	1	$[-\infty, \infty]$
gammasShape.s:noncoding	Exponential	1	$[-\infty, \infty]$
Kappa.s:1stpos	Log normal	2	$[0, 0, \infty]$
Kappa.s:2ndpos	Log normal	2	$[0, 0, \infty]$
Kappa.s:3rdpos	Log normal	2	$[0, 0, \infty]$
Kappa.s:coding	Log normal	2	$[0, 0, \infty]$
Kappa.s:noncoding	Log normal	2	$[0, 0, \infty]$

Table S8. Priors applied for the different tests of the coalescence constant and exponential population growth.

Parameter	Distribution	Initial value	Interval
ePopSize.t.tree	Gamma	1	$[0, 001, 1000]$
ClockRate.c:clock	Uniform	1	$[-\infty, \infty]$
uclMean.c:clock (PDVall)	Uniform	0.0014	$[0, 0.1]$
uclMean.c:clock (1988 and 2002)	Uniform	4.329E-4	$[0, 0.1]$
GrowthRate.t.tree	Laplace Distribution	3.0E-4	$[-\infty, \infty]$

Table S9. Mutations found in the sequences from samples collected in 1988. The parentheses indicate the positions where a fragment from a German sample was used to fill gaps. Results from an additional sample (42523) which was not fully sequenced was included and non-sequenced areas marked with -.

Samples	Locality	1.868	2.251	2.481	3.131	3.357	4.562	4.869	4.978	5.126	5.202	5.816	5.891	7.414	7.444	7.602	8.023	8.201
		(Q = 52)	(Q = 52)	(Q = 55)	(Q = 55)	(Q = 52)	(Q = 52)	(Q = 54)	(Q = 59)	(Q = 30)							(Q = 52)	
PV14965	Germany	C	A	C	G	G	G	A	C	C	C	T	C	T	A	T	C	A
PV14958	Germany	C	A	C	G	G	G	A	C	C	C	T	C	T	A	T	A	A
PV14948	Germany	C	A	C	G	A	G	A	C	C	C	T	C	T	A	T	C	A
PV14969	Germany	C	G	C	G	G	A	A	C	C	C	T	C	T	A	T	A	A
42524	Limfjorden	C	A	T	A	G	G	A	C	C	T	G	T	C	G	T	A	A
42523	Limfjorden	-	-	-	-	-	-	-	-	C	T	G	T	C	G	T	A	A
42531	Limfjorden	C	A	C	G	G	G	C	C	T	C	T	C	T	A	T	A	A
42412	Anholt	A	A	C	G	(G)	(G)	A	C	C	C	T	T	T	A	G	A	A
42413	Anholt	C	A	C	G	(G)	(G)	A	T	C	C	T	T	T	A	T	A	G

Table S10. The mutations (≥ 2) of the PDV2002 data set. Samples are listed with a number indicating the clade to which they belong to in the tree.

Sample	Area	1.792 (P gene)	1.795 (P gene)	2.152 (P gene)	2.473 (P gene)	3.614 (M gene)	4,010 (M gene)	4.698 (Non c. M gene)	4.747 (Non c. M gene)	5.064 (F gene)	5.266 (F gene)	5.306 (F gene)	6.893 (F gene)	7.129 (H gene)	7.146 (H gene)	7.447 (H gene)
0_25002	DK-KAT	T	T	G	T	T	C	T	A	C	T	A	G	G	T	C
1_25030 ⁴	DK-KAT	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
1_30030 ⁴	DK-KAT	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
1_30031 ⁴	DK-KAT	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
1_VB21 ³	SE-KAT	T	T	G	C	T	C	C	A	C	T	A	G	G	T	C
1_VB22 ³	SE-KAT	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
1_TM3155	SE-KAT	T	T	G	C	T	C	C	A	C	T	A	G	G	T	C
1_TM3107	SE-SKG	T	C	G	C	T	C	T	A	C	T	A	G	G	T	C
1_TM3118	SE-SKG	T	C	G	C	T	C	T	A	T	T	A	G	G	T	C
1_TM3192	SE-SKG	T	C	G	C	T	C	T	A	T	T	A	G	G	T	C
1_TM16024	SE-SKG	T	C	G	C	T	C	T	A	T	T	A	G	G	T	C
1_26055	DK-ØRE	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
1_25677	DK-BAL	T	T	G	C	T	C	T	A	C	T	A	G	G	T	C
2_26866	DK-LIM	T	T	G	T	T	C	T	A	C	T	A	G	A	T	C
2_26891	DK-LIM	T	T	G	T	T	C	T	A	C	T	A	G	A	T	C
2_26901	DK-LIM	T	T	G	T	T	C	T	A	C	T	A	G	A	T	C
2_30029	DK-KAT	T	T	G	T	T	C	T	A	C	T	A	G	A	T	C
3_EPDV2681	UK-Wash	T	T	G	T	C	C	T	A	C	T	A	G	G	T	C
3_EPDV2480	UK-Wash	T	T	G	T	C	C	T	A	C	T	A	A	G	T	C
3_EPDV1909	UK-Wash	T	T	G	T	C	C	T	A	C	T	A	A	G	T	C
3_EPDV1744	UK-Wash	T	T	G	T	C	C	T	A	C	T	A	A	G	T	C
4_EPDV1966	UK-Wash	T	T	G	T	T	T	T	G	C	T	A	G	G	T	C
4_H41	NL-WAD	T	T	G	T	T	T	T	A	C	T	A	G	G	T	A
4_Pv 1939	DE-WAD	C	T	G	T	T	T	T	A	C	G	A	G	G	T	A
4_Pv 2149	DE-WAD	T	T	G	T	T	T	T	G	C	T	A	G	G	T	A
4_Pv 2162	DE-WAD	T	T	G	T	T	T	T	A	C	T	A	G	G	T	A
4_Pv 2163 ²	DE-WAD	T	T	G	T	T	T	T	A	C	T	A	G	G	T	A
4_26001 ²	DK-WAD	T	T	G	T	T	T	T	A	C	T	A	G	G	T	A
4_26002 ¹	DK-WAD	C	T	A	T	T	T	T	A	C	G	G	G	G	T	A
4_26003	DK-WAD	C	T	G	T	T	T	T	A	C	G	A	G	G	T	A
4_26561 ¹	DK-WAD	C	T	A	T	T	T	T	A	C	G	G	G	G	T	A
5_VB23	SE-KAT	T	T	G	T	T	C	T	A	C	T	A	G	G	T	C
5_TM197	SE-SKG	T	T	G	T	T	C	T	A	T	T	A	G	G	C	C
5_TM3109	SE-SKG	T	T	G	T	T	C	T	A	C	T	A	G	G	C	C
5_25679	DK-BAL	T	T	G	T	T	C	T	A	C	T	A	G	G	T	C
5_25682	DK-BAL	T	T	G	T	T	C	T	A	C	T	A	G	G	T	C