Cell Reports, Volume *26*

Supplemental Information

The N-Terminal Domain of cGAS Determines Preferential Association with Centromeric DNA and Innate Immune Activation in the Nucleus

Matteo Gentili, Xavier Lahaye, Francesca Nadalin, Guilherme F.P. Nader, Emilia Puig Lombardi, Solène Herve, Nilushi S. De Silva, Derek C. Rookhuizen, Elina Zueva, Christel Goudot, Mathieu Maurin, Aurore Bochnakian, Sebastian Amigorena, Matthieu Piel, Daniele Fachinetti, Arturo Londoño-Vallejo, and Nicolas Manel

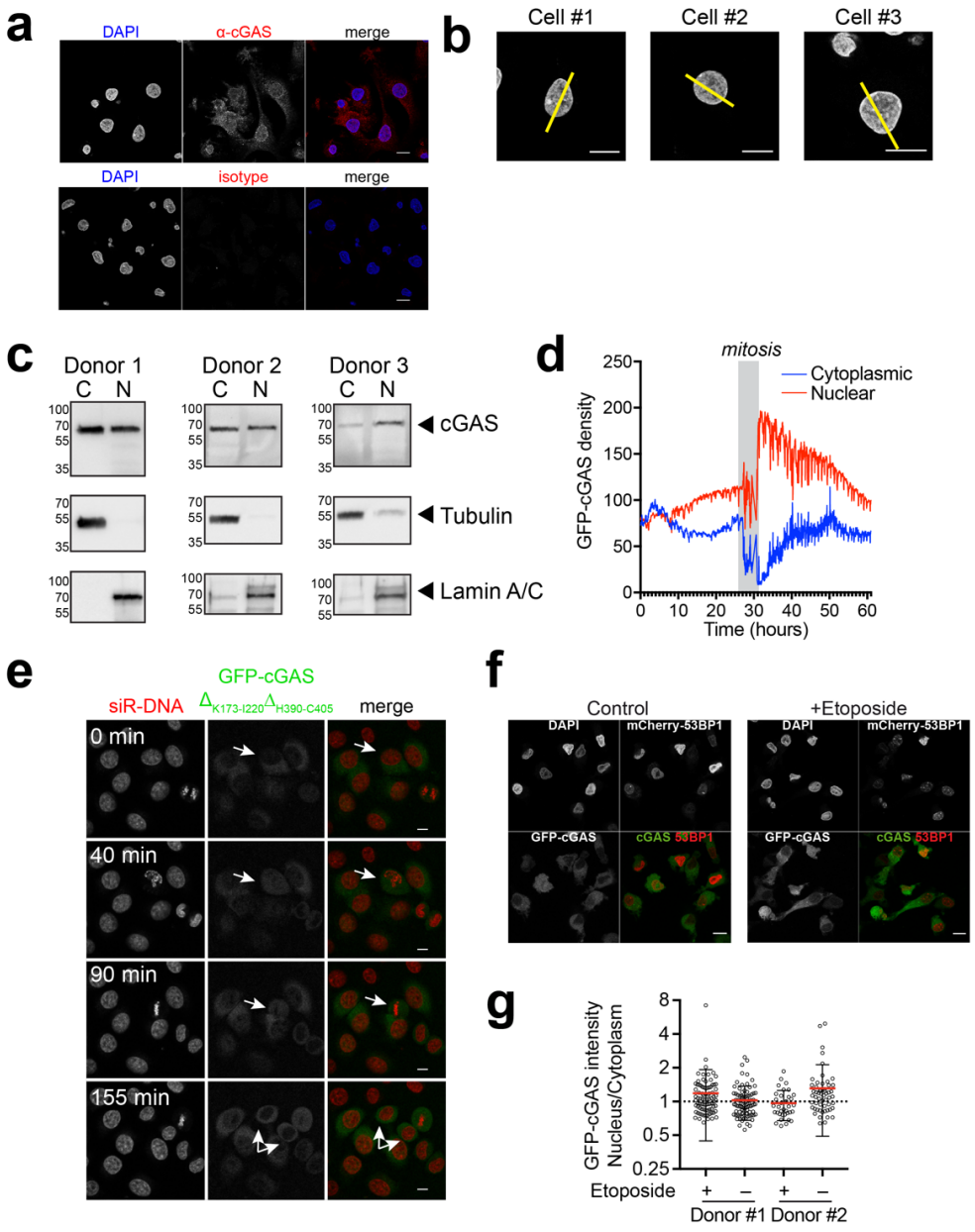


Figure S1. cGAS is present in the nucleus as a result of nuclear envelope opening and DNA binding, related to Figure 1. (a) Immunofluorescence staining of endogenous cGAS or isotype control (red) and DAPI (blue) in post-mitotic human monocyte-derived dendritic cells (DCs). One field for one donor representative of n=4 donors. Scale bar is 10µm. **(b)** Magnification of the DAPI channel for cells in Figure 1a. Scale bars are 10µm. **(c)** Nuclear/Cytoplasmic fractionation of DCs and immunoblot for cGAS (top), the cytosolic marker Tubulin (middle) and the nuclear marker Lamin B1 for Donor 2 or Lamin A/C for Donor 3 and 4 (bottom), related to Figure 1c. **(d)** Quantification of nuclear GFP-cGAS intensity in the nucleus during mitosis and the following interphase, related to Movie S3. **(e)** Sequential images of HeLa cells expressing GFP-cGAS $\Delta_{K173\text{-}I220}\Delta_{H390\text{-}405}$ (green) in which the DNA has been stained with siR-DNA (red). Arrows indicate a dividing cell. The two arrows at 155 minutes indicate the two daughter cells from the initial one. **(f)** Immunofluorescence of DCs transduced with mCherry-53BP1$_{1224\text{-}1716}$ in pTRIP-CMV (red) and GFP-cGAS in pTRIP-SFFV (green) and untreated (left) or treated (right) with 50µM Etoposide for 24h. One donor representative of two independent donors. Scale bars are 10µm. **(g)** Quantification of nuclear/cytoplasmic ratio of GFP-cGAS in DCs treated as in (f). Red line=mean, error bar=standard deviation. Each dot represents a single cell. n≥34 cells for each condition. n=2 independent donors.

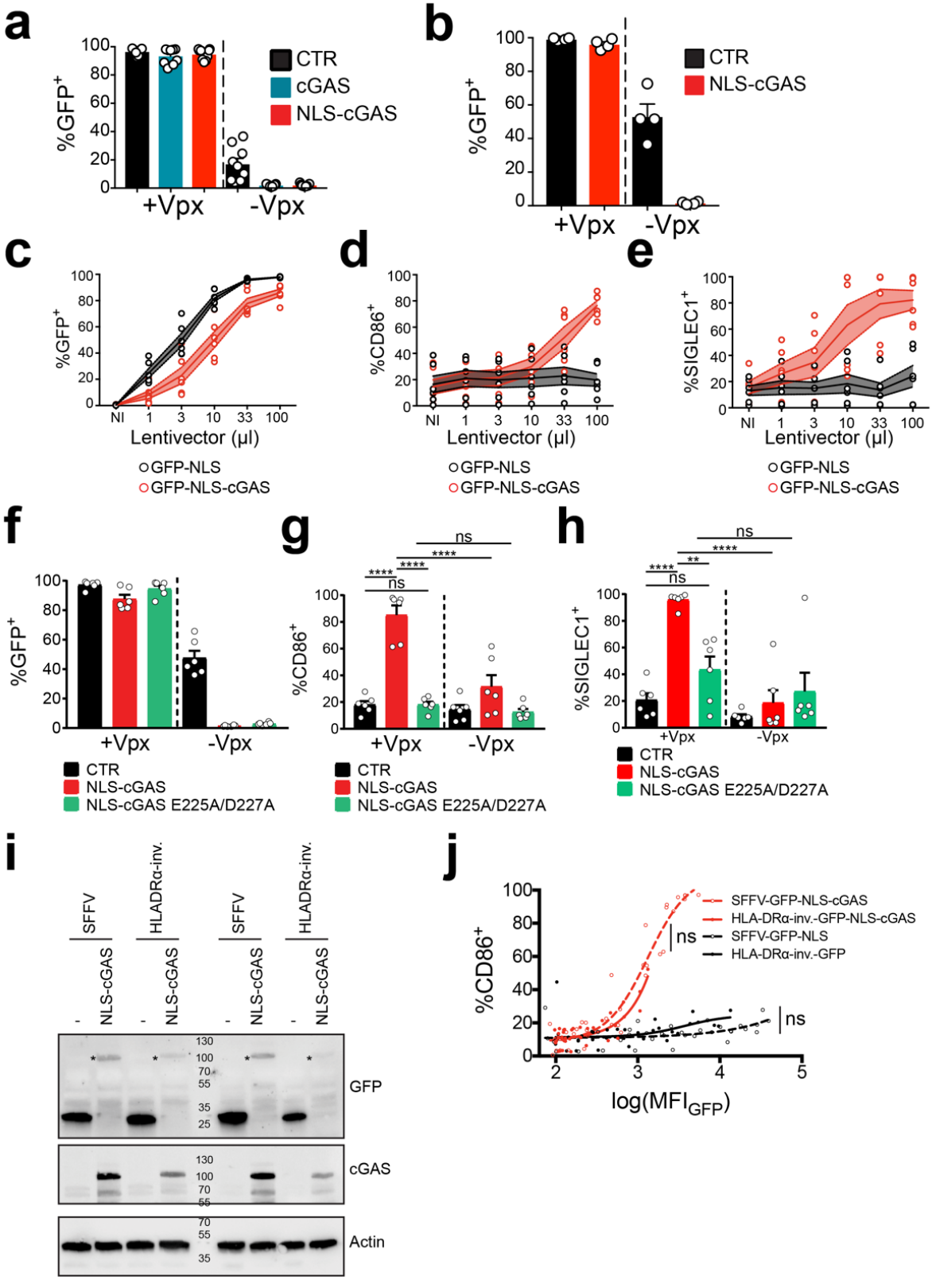

a
%GFP+
CTR
cGAS
NLS-cGAS
+Vpx
-Vpx
b
%GFP+
CTR
NLS-cGAS
+Vpx
-Vpx
c
%GFP+
NI
1
3
10
33
100
Lentivector (µl)
GFP-NLS
GFP-NLS-cGAS
d
%CD86+
Lentivector (µl)
GFP-NLS
GFP-NLS-cGAS
e
%SIGLEC1+
Lentivector (µl)
GFP-NLS
GFP-NLS-cGAS
f
%GFP+
+Vpx
-Vpx
CTR
NLS-cGAS
NLS-cGAS E225A/D227A
g
%CD86+
ns
****
****
****
ns
+Vpx
-Vpx
CTR
NLS-cGAS
NLS-cGAS E225A/D227A
h
%SIGLEC1+
ns
****
**
****
ns
+Vpx
-Vpx
CTR
NLS-cGAS
NLS-cGAS E225A/D227A
i
SFFV
HLADRα-inv.
SFFV
HLADRα-inv.
-
NLS-cGAS
-
NLS-cGAS
-
NLS-cGAS
-
NLS-cGAS
130
100
70
55
35
25
GFP
cGAS
Actin
j
%CD86+
SFFV-GFP-NLS-cGAS
HLA-DRα-inv.-GFP-NLS-cGAS
SFFV-GFP-NLS
HLA-DRα-inv.-GFP
ns
ns
log(MFIGFP)

Figure S2. Nuclear-localized cGAS activates an innate immune response in dendritic cells, related to Figure 2. (a) GFP expression in DCs related to transductions of Figure 2c, 2d. Each dot represents an individual donor. n=9 donors in four independent experiments. **(b)** GFP expression in DCs related to transductions of Figure 2f, 2g, 2h, 2i. Each dot represents an individual donor. n=6 donors in three independent experiments. **(c)** GFP expression in DCs following dose titration of GFP-NLS or GFP-NLS-cGAS lentivectors in pTRIP-SFFV with Vpx, related to Figure 2f, 2g, 2h, 2i. Each dot represents an individual donor. Solid lines represent the mean of the experiments and the lighter colored limits represents SEM. n=6 donors in three independent experiments. **(d)** CD86 expression in dose titration as in (c). Each dot represents an individual donor. Solid lines represent the mean of the experiments and the lighter colored limits represents SEM. n=6 donors in three independent experiments. **(e)** SIGLEC1 expression in dose titration as in (c). Each dot represents an individual donor. Solid lines represent the mean of the experiments and the lighter colored limits represents SEM. n=6 donors in three independent experiments. **(f)** GFP expression in DCs transduced with GFP-NLS, GFP-NLS-cGAS or catalytically dead GFP-NLS-cGAS E225A/D227A lentivectors in pTRIP-SFFV, in presence or in absence of Vpx. Each dot represents an individual donor. n=6 donors of three independent experiments. **(g)** CD86 expression in DCs transduced as in (f). Each dot represents an individual donor. n=6 donors of three independent experiments. One-way ANOVA with post-hoc Tukey test; ****P<0.0001, ns=non-significant. **(h)** SIGLEC1 expression in DCs transduced as in (f). One-way ANOVA with post-hoc Tukey test; ****P<0.0001, **P<0.01, ns=non-significant. **(i)** Immunoblot of GFP, cGAS and actin in DCs expressing either GFP or GFP-NLS-cGAS (*) under the control of an SFFV or an HLA-DRα inverted promoter. n=2 donors. One donor is at the left of the ladder, the other donor is at the right of the ladder. **(j)** Correlation between Mean Fluorescence Intensity (MFI) of GFP and %CD86 expression in DCs transduced with a control GFP vector (black) or a GFP-NLS-cGAS vector (red) under the control of a SFFV promoter (dashed lines) or an HLA-DRα inverted promoter (solid lines). Each dot represents an individual donor. Lines are interpolated using a four-parameter dose-response curve equation. Kruskal-Wallis with Dunn's multiple comparisons test; ns=non-significant.

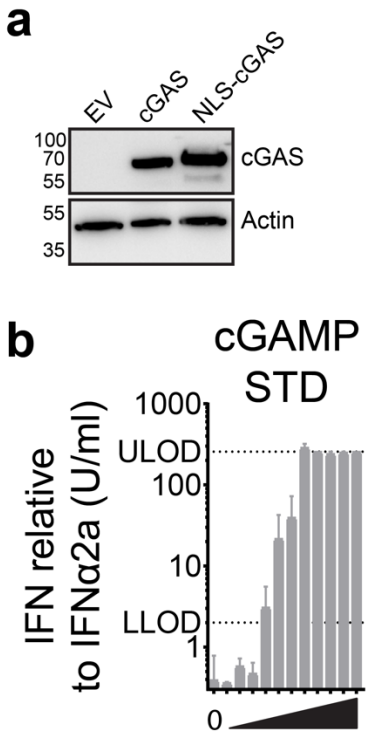
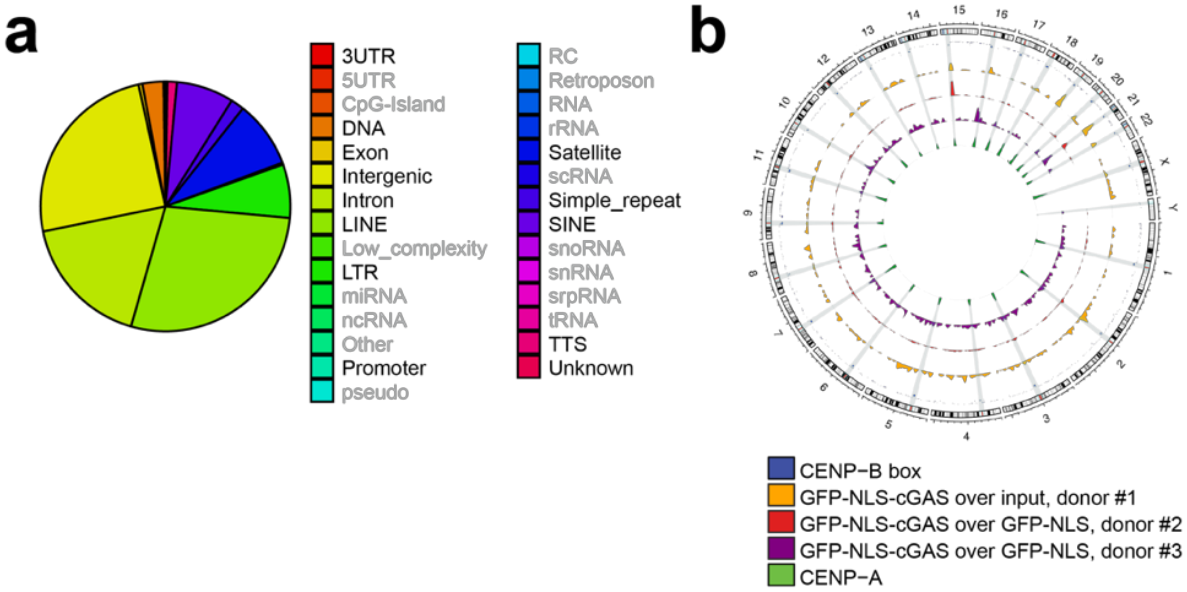
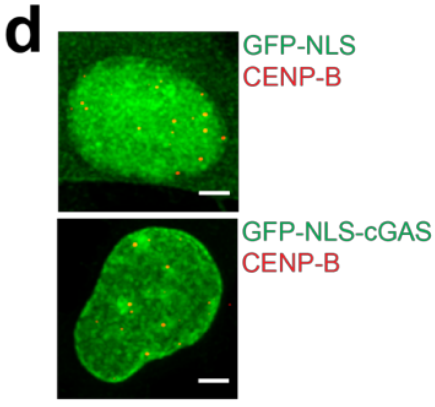


Figure S3. Nuclear localization of cGAS results in limited cGAMP production, related to Figure 3. (a) Immunoblot of cGAS (top) and Actin (bottom) of HEK293FT stably transduced with a control vector (EV), cGAS or NLS-cGAS in pTRIP-CMV. **(b)** Type I IFN activity measured on HL-116 cells of supernatant coming from PMA-differentiated and permeabilized THP-1 stimulated with synthetic 2'3'-cGAMP. Mean and SEM of n=3 independent experiments. Top dose is 156ng. Dilutions are 3-fold.



c

Motif	K-mer sig	E-value
	53.22	6E-54
	53.22	6E-54
	53.22	6E-54
	53.22	6E-54
	3.47	0.00034



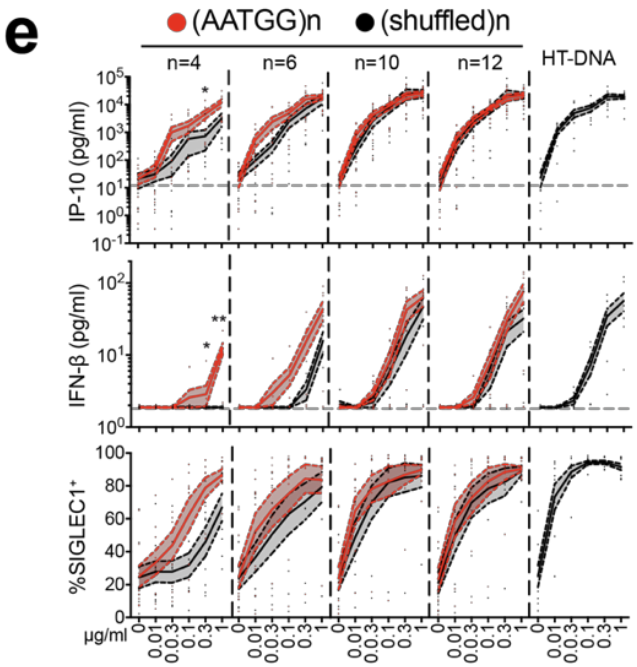


Figure S4. Nuclear-localized cGAS associates with centromeric DNA, related to Figure 5. (a) Distribution of GFP-cGAS annotated filtered peaks (n=404 peaks) over input across genomic elements (1 donor representative of 3 independent donors). Elements with less than 10 peaks are grayed out. **(b)** Distribution of GFP-NLS-cGAS and of CENP-A peaks and localization of CENP-B box (consensus sequence) on the hg38 genome. The cGAS-NLS-cGAS tracks represents the density of filtered cGAS peaks (over input for donor #1, over GFP-NLS ChIP for donors #2 and #3) and is computed on windows of size 10^7 across the genome. 404, 754, and 762 filtered peaks are identified for the three donors, respectively. The CENP-A track represents the density of CENP-A intersection peaks and is computed on windows of size 10^7 across the genome. The CENP-B box track reports on the x axis the genomic position of the region (occurrence of CENP-B box consensus sequence) and on the y axis the minimal distance (log10 transformed) of the region to its two neighbouring regions. Centromere locations and their flanking regions (5Mbp upstream and downstream) are highlighted in cyan. Centromere location are retrieved from cytobands coordinates (UCSC). **(c)** *De novo* motif discovery in the sequences of cGAS intersection peaks (GFP-NLS-cGAS over GFP-NLS ChIP-seq) between donor #2 and donor #3. **(d)** Maximum intensity Z projection of the Z-stacks shown in Figure 5h and Figure 5j. **(e)** IP-10, IFN-ß and SIGLEC1 expression by DCs stimulated by transfected synthetic DNA repeats coding for the AATGG satellite motif or the shuffled sequence, or HT-DNA, at the indicated DNA concentrations (independent donors: n=9 for IP-10 and SIGLEC1, n=7 for IFN-ß; two-way ANOVA with Tukey test, on log-transformed data for IP-10 and IFN-ß; **P<0.01, *P<0.05).

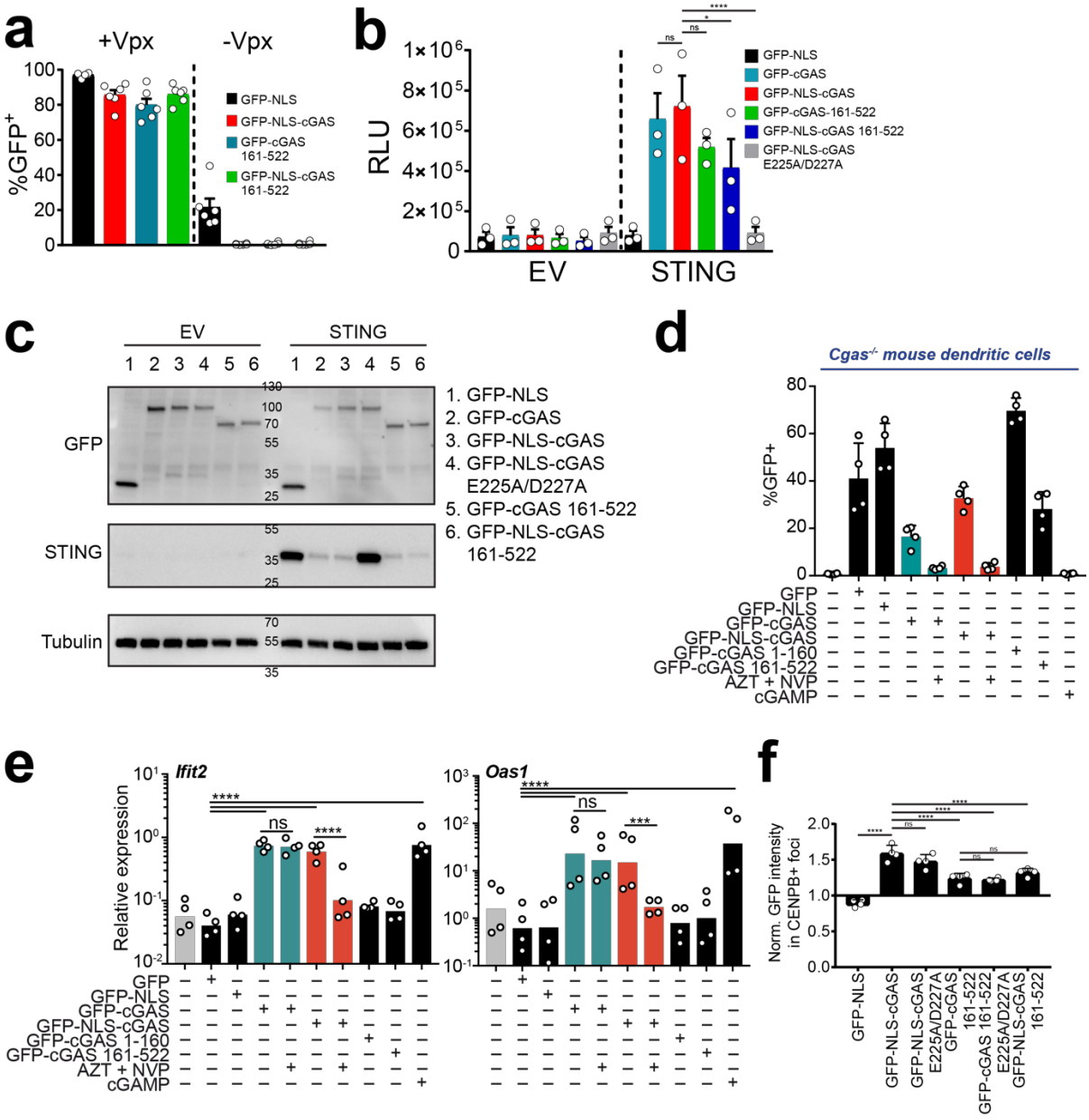


Figure S5. The N-terminal domain of cGAS mediates association with centromeric DNA, related to Figure 6. (a) GFP expression in DCs related to transductions of Figure 6c, 6d. Each dot represents an individual donor. n=6 donors in three independent experiments. **(b)** Quantification of Relative Light Units (RLU) from 293FT cells co-transfected with a vector encoding for Firefly Luciferase under the control of the IFNß promoter, an Empty Vector (EV) or a vector encoding for human STING, in presence of a control vector (GFP-NLS) or of the indicated GFP-cGAS constructs in pTRIP-SFFV. One-way ANOVA with post-hoc Tukey test. ****P<0.001, *P<0.05, ns=non-significant. **(c)** Immunoblot of GFP (top), STING (middle) and α-Tubulin (bottom) of 293FT cells transfected as in (b). **(d)** Expression of GFP in *Cgas*$^{-/-}$ mouse bone marrow-derived DCs transduced with GFP, GFP-NLS, GFP-cGAS, GFP-NLS-cGAS, GFP-cGAS 1-160 or GFP-cGAS 161-522 in pTRIP-SFFV lentivectors, untreated or treated with reverse transcriptase inhibitors (AZT + NVP), or transfected with cGAMP (n=4 mice combined from 2 independent experiments). **(e)** Expression of *Ifit2* and *Oas1* in cells as in Figure 6f, bars represent geometric mean (One-way ANOVA with Sidak test, on log-transformed data; ****P<0.0001, *** P<0.001, ns=non-significant). **(f)** Mean GFP intensity in CENP-B foci normalized to mean nuclear GFP intensity for GFP-NLS and for the indicated GFP-cGAS constructs for n=4 donors. Each dot represents an individual donor. One-way ANOVA with post-hoc Tukey test; ****P<0.0001, ns=non-significant.

Table S1. Oligonucleotides used in the study, related to Key Resource Table.

NAME	SEQUENCE	SOURCE
mRPS29-qPCRB-f	GAGCCGACTCGTTCCTTT	Eurogentec
mRPS29-qPCRB-r	TGTTCAGCCCGTATTTGC	Eurogentec
ms_Ifit1-0f	CAAGGCAGGTTTCTGAGGAG	Eurogentec
ms_Ifit1-98r	GACCTGGTCACCATCAGCAT	Eurogentec
ms_Ifit2-5f	AAGCAAGTTCTGGCCTTCTG	Eurogentec
ms_Ifit2-111r	AGCAGCTGGTTCCTTTTCCT	Eurogentec
ms_Oas1a-99f	CTGCATCAGGAGGTGGAGTT	Eurogentec
ms_Oas1a-195r	GGATGGCATAGATTCTGGGA	Eurogentec
bactin737f	GGACTTCGAGCAAGAGATGG	Eurogentec
bactin970r	AGCACTGTGTTGGCGTACAG	Eurogentec
mx1-662-f	TTACCAGGACTACGAGATTGAG	Eurogentec
mx1-893-r	GATGAGTGTCTTGATCTTATACCC	Eurogentec
oas1-f	GAGCTCCAGGGCATACTGAG	Eurogentec
oas1-r	CCAAGCTCAAGAGCCTCATC	Eurogentec
IFIT1-143f	CAACCATGAGTACAAATGGTG	Eurogentec
IFIT1-425r	TGGCATTCAAGGAGTACCTC	Eurogentec
CXCL10-116f	TGGCATTCAAGGAGTACCTC	Eurogentec
CXCL10-261r	TTGTAGCAATGATCTCAACACG	Eurogentec
NM_(sat)4-f	AATGGAATGGAATGGAATGG	Eurogentec
NM_(sat)4-r	CCATTCCATTCCATTCCATT	Eurogentec
NM_(sat)6-f	AATGGAATGGAATGGAATGGAATGGAATGG	Eurogentec
NM_(sat)6-r	CCATTCCATTCCATTCCATTCCATTCCATT	Eurogentec
NM_(sat)10-f	AATGGAATGGAATGGAATGGAATGGAATG GAATGGAATGGAATGGAATGG	Eurogentec
NM_(sat)10-r	CCATTCCATTCCATTCCATTCCATTCCATTCC ATTCCATTCCATTCCATT	Eurogentec
NM_(sat)12-f	AATGGAATGGAATGGAATGGAATGGAATGG AATGGAATGGAATGGAATGGAATGGAATGG	Eurogentec
NM_(sat)12-r	CCATTCCATTCCATTCCATTCCATTCCATTCC ATTCCATTCCATTCCATTCCATTCCATT	Eurogentec
NM_(shuf)4-f	GAGAGTTGATAGTGGAGAAA	Eurogentec
NM_(shuf)4-r	TTTCTCCACTATCAACTCTC	Eurogentec
NM_(shuf)6-f	AAGTTAAGGGAAGAGTGATGTGGGAAGTAA	Eurogentec
NM_(shuf)6-r	TTACTTCCCACATCACTCTTCCCTTAACTT	Eurogentec
NM_(shuf)10-f	GGAAAAGGATTGGGGATGGGGGATTGGTT AAAATGTGGAAGAAAAAGTAA	Eurogentec
NM_(shuf)10-r	TTACTTTTTCTTCCACATTTTAACCAATCCCC CATCCCCAATCCTTTTCC	Eurogentec
NM_(shuf)12-f	AGGAGTTAATTATGGTATGGTATGGAAGAAA AAAAAAGGGGAGATGGGTGGAAGAAGGTG	Eurogentec
NM_(shuf)12-r	CACCTTCTTCCACCCATCTCCCCTTTTTTTTTC TTCCATACCATACCATAATTAACTCCT	Eurogentec