Ladder-like multimerization of FoxP3 enables microsatellite recognition and DNA bridging

Wenxiang Zhang[1,2,5], Fangwei Leng[1,2,5], Xi Wang[1,2], Ricardo N. Ramirez[3,4], Jinseok Park[3], Christophe Benoist[3,4], and Sun Hur[1,2]*

[1]Howard Hughes Medical Institute and Program in Cellular and Molecular Medicine, Boston Children's Hospital, MA 02115, USA

[2]Department of Biological Chemistry and Molecular Pharmacology

[3]Department of Immunology, Harvard Medical School, Boston, MA 02115, USA

[4]Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA 02115, USA

[5]These authors contributed equally.

*Correspondence: Sun.Hur@crystal.harvard.edu

Abstract:
FoxP3 is a transcription factor (TF) essential for development of regulatory T cells (Tregs), a branch of T cells that suppress excessive inflammation and autoimmunity[1-5]. Molecular mechanisms of FoxP3, however, remain elusive. We here show that FoxP3 utilizes the Forkhead domain—a DNA binding domain (DBD) that is commonly thought to function as a monomer or dimer—to form a higher-order multimer upon binding to T_nG repeat microsatellites. A cryo-electron microscopy structure of FoxP3 in complex with T_3G repeats reveals a ladder-like architecture, where two double-stranded DNA molecules form the two "side rails" bridged by five pairs of FoxP3 molecules forming "rungs". Each FoxP3 subunit occupies TGTTTGT within the repeats in the manner indistinguishable from that of FoxP3 bound to the Forkhead consensus motif (FKHM; TGTTTAC). Mutations in the "intra-rung" interface impair T_nG repeat recognition, DNA bridging and cellular functions of FoxP3, all without affecting FKHM binding. FoxP3 can tolerate variable "inter-rung" spacings, explaining its broad specificity for T_nG repeat-like sequences *in vivo* and *in vitro*. Both FoxP3 orthologs and paralogs show similar T_nG repeat recognition and DNA bridging. These findings thus reveal a new mode of DNA recognition that involves TF homo-multimerization and DNA bridging, and further implicates microsatellites in transcriptional regulation and diseases.

Introduction

How TFs utilize a limited repertoire of DBDs to orchestrate complex gene regulatory networks is a central and yet unresolved question[6-9]. While certain TFs, such as those with zinc-finger (ZF) DBDs, can expand the complexity of their sequence specificity by forming an array of DBDs, the vast majority of TFs utilize a single DBD with narrow sequence specificity shared with other members of the DBD family[7]. One prominent model to rationalize this apparent paradox is that cooperative actions of multiple distinct TFs with distinct DBDs give rise to combinatorial complexity[10,11]. However, whether a single TF with a single DBD can also recognize distinct sequences on its own and perform divergent transcriptional functions, depending on the conformation or multimerization state, has not been fully addressed.

FoxP3 is an essential TF in Treg development, of which loss-of-function mutations cause a severe multiorgan autoimmune disease, immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome[1-5]. Previous studies showed that FoxP3 remodels the global transcriptome and three-dimensional genome organization in the late stage of Treg development[12-15]. However, molecular mechanisms of FoxP3, including its direct target genes and *in vivo* sequence specificity, remain unclear[13-16].

FoxP3 DNA binding is primarily mediated by a Forkhead domain, which is shared among ~50 TFs of the Forkhead family[17,18]. Most Forkhead domains form a conserved winged-helix conformation and recognize the consensus FKHM sequence (TGTTTAC)[19]. While isolated Forkhead domain of FoxP3 was originally crystallized as an unusual domain-swap dimer[20,21], a recent study showed that FoxP3 does not form a domain-swap dimer but instead folds into the canonical winged-helix conformation in the presence of the adjacent Runx1-binding region (RBR)[22]. It was further shown that FoxP3 has a strong preference for inverted repeat FKHM (IR-FKHM) over a single FKHM in vitro by forming a head-to-head dimer[22]. However, previous chromatin immunoprecipitation sequencing (ChIP-seq)[14,23,24] and Cleavage Under Targets & Release Using Nuclease sequencing (CNR-seq)[14] analyses did not reveal enrichment of IR-FKHM in FoxP3-occupied genomic regions within cells[22]. While individual FKHM is present in ~10% of the FoxP3 ChIP peaks, they too may not be the FoxP3 binding sites as DNase I protection patterns at these sites were unaffected by FoxP3 deletion[24]. These observations raised the question of what sequences FoxP3 in fact recognizes in cells and whether FoxP3 can employ a previously unknown mode of binding to recognize new sequence motifs distinct from FKHM.

Results

FoxP3 recognizes T_nG repeat microsatellite *in vitro* and *in vivo*.

To re-evaluate FoxP3 sequence specificity, we performed an unbiased pull-down of genomic DNA with recombinant FoxP3 protein. Use of genomic DNA, as opposed to synthetic DNA oligos, enables testing sequence specificity in the context of naturally existing repertoire of sequences. It can also allow identification of longer motifs by using genomic DNA fragments longer than ~20-40 bp—the typical lengths used in previous biochemical studies of FoxP3[22,25,26].

We isolated genomic DNA from mouse EL4 cells, fragmented to ~100-200 bp, incubated with purified, MBP-tagged mouse FoxP3 and performed MBP pull-down, followed by next-generation sequencing (NGS) of the co-purified DNA (FoxP3 PD-seq, Fig. 1a). We used recombinant FoxP3 protein ($FoxP3^{\Delta N}$) harboring zinc finger (ZF), coiled coil (CC), RBR and Forkhead domains but lacking the N-terminal Pro-rich region (Fig. 1a). $FoxP3^{\Delta N}$ was previously shown to display the same DNA specificity as full-length FoxP3 among the test set[22]. *De novo* motif analysis showed a strong enrichment of T_nG repeats (n=2-5) by FoxP3 pull-down, using either MBP alone pull-down or input as a control (Fig. 1b, Extended Data Fig. 1a). T_3G repeat sequence was the highest-ranking motif, which accounted for 40% of the peaks. No other motifs, including the canonical FKHM or other repeats, were similarly enriched (Extended Data Fig. 1a).

Re-analysis of previously published FoxP3 CNR-seq and ChIP-seq data[14] using their consensus peaks (CNR/ChIP peaks) identified T_nG repeat-like motifs in 23.5% of the CNR/ChIP peaks (E-value of 4.7e-009) (Fig. 1b). The enrichment score for T_nG repeat-like motifs was more significant than that of FKHM (10.8% with E-value of 1.5e+000). A similar T_nG repeat-like motif was also observed with independent FoxP3 ChIP-seq data[23,24] (Extended Data Fig. 1b). T_nG repeat-like motifs (n>1) were not identified from open chromatin regions (as measured by ATAC-seq[27]) in Treg cells that were not occupied by FoxP3 (Extended Data Fig. 1b). Genomic feature analysis of the CNR/ChIP peaks with and without T_nG repeat-like sequences suggested that T_nG-containing peaks were enriched for TSS-distal regions, while T_nG-depleted peaks were enriched for TSS-proximal regions, such as promoters (Extended Data Figs. 1c, 1d). Consistent with this, ~82% of the T_nG-depleted peaks showed a high level of trimethylated H3K4 (H3K4me3 > 0.2), the histone mark indicative of promoters, while 48% of the T_nG-containing peaks showed H3K4me3 > 0.2 (Extended Data Fig 1e).

To examine whether T_nG repeat-like sequences indeed contribute to FoxP3–DNA interaction in Treg cells, we analyzed published FoxP3 CNR-seq data generated using F1 hybrids of B6 and Cast mice strains[14]. Because of the wide divergence between B6 and Cast genomes, such data allow evaluation of the impact of sequence variations on TF binding. Out of 196 sites showing allelic imbalance (fold change >=4) in FoxP3 CNR-seq, 78 sites harbored T_nG repeat-like elements in at least one allele. Of the 78 sites, we randomly chose 50 sites, 25 each from B6- and Cast-biased peaks, and tested FoxP3 binding efficiency by $FoxP3^{\Delta N}$ pull-down. Out of the 50 pairs of sequences tested, pull-down efficiency of 47 pairs recapitulated differential binding in CNR-seq (Figs. 1c, 1d). All 47 sites showed significant truncations in the T_nG repeats in the less-preferred allele (Fig. 1c; see Extended Data Table 1 for the full list and sequences). Note that the pull-down preference for longer T_nG repeats was not due to the different DNA lengths used—an extension of the less-preferred allele sequences with a random sequence at a DNA end (B6* and Cast* in Fig. 1c, see Extended Data Table 2 for sequence) did not alter the allele bias. Altogether, these results suggest that T_nG repeat-like elements play an important role in FoxP3–DNA interaction *in vitro* and *in vivo*.

T_nG repeat recognition is mediated by FoxP3 homo-multimerization.

To examine T_nG repeat enrichment in PD-seq and CNR/ChIP-seq indeed represent previously unrecognized sequence specificity of FoxP3, we compared FoxP3 binding to DNA with T_nG repeats (n=1 to 6) vs. those containing IR-FKHM, the highest affinity sequence reported for FoxP3 to date[22]. All DNAs were 45 bp-long (see Extended Data Table 2 for sequences). We found that T_3G repeat was comparable to IR-FKHM in FoxP3 binding and was the tightest binder among the T_nG repeats (Fig. 1e), consistent with it being the most significant motif in PD-seq (Fig. 1b). T_2G, T_4G and T_5G repeats also showed more efficient binding than a single FKHM (1XFKHM) or random sequence (no FKHM). No other simple repeats showed FoxP3 binding comparable to T_3G repeats (Fig. 1f). FoxP3 affinity increased with the copy number of T_3G, when compared among DNAs of the same length (Fig. 1g). The preference for T_3G repeats was also seen using full-length FoxP3 expressed in HEK293T cells or when pull-down bait was switched from FoxP3 to DNA (Extended Data Fig. 2a, 2b).

We next investigated how FoxP3 recognizes T_3G repeats. Unlike IR-FKHM, T_3G-repeat DNA induced FoxP3 multimerization as indicated by slowly migrating species in native gel-shift assay (Fig. 1h). Protein-protein crosslinking also suggested higher-order multimerization in the presence of T_3G repeats, but not with IR-FKHM or 1XFKHM (Extended Data Fig. 2c). In support of T_3G repeat-induced multimerization, MBP-tagged FoxP3 was co-purified with GST-tagged FoxP3 only in the presence of T_3G repeat, but not with IR-FKHM (Extended Data Fig. 2d). Finally, negative electron microscopy revealed a filamentous multimeric architecture of FoxP3 on 36 tandem repeats of T_3G (Fig. 1i), the copy number chosen to aid clear visualization. Other DNAs of the same length, such as $(A_3G)_{36}$ or $(\text{IR-FKHM})_6$, did not show similar multimeric architectures under equivalent conditions (Extended Data Fig. 2e). These results suggest that FoxP3 forms distinct multimers on T_3G repeats.

Cryo-EM structure reveals a ladder-like assembly of FoxP3 on T_nG repeat DNA.

To understand how FoxP3 forms multimers on T_3G repeats, we determined the cryo-electron microscopy (cryo-EM) structure of $FoxP3^{\Delta N}$ in complex with $(T_3G)_{18}$. Single particle reconstruction led to a 3.7 Å resolution map after global refinement and a 3.3 Å resolution map after focused refinement of the central region (Extended Data Fig. 3). The density map revealed two continuous double-stranded DNA molecules spanning ~50 bp (Fig. 2a). Both DNA molecules adopted the classic B-form DNA with the average twist angle of 33.5˚/bp and the average rise of 3.19 Å/bp. The density map could also be fitted with the crystal structure of DNA-bound FoxP3 monomer containing part of RBR and Forkhead (residue 326-412), allowing placement of ten FoxP3 subunits without ZF, CC and RBR residues 262-325. Only the non-swap, winged-helix conformation was compatible with the density map (Extended Data Fig. 4a). Consistent with this, R337Q, a loss-of-function IPEX mutation that induces domain-swap dimerization[22], showed significantly reduced affinity for T_3G repeats (Extended Data Fig. 4b).

DNA sequence assignment (see Extended Data Fig. 5a) revealed that all ten FoxP3 subunits interacted with T_3G repeat DNA in the manner indistinguishable from that of FoxP3 bound to the canonical FKHM, recognizing TGTTTGT in place of TGTTTAC (Fig. 2b). This FoxP3-DNA register was further confirmed by FoxP3 footprint analysis using DNA mutagenesis and NFAT–

FoxP3 cooperativity (Extended Data Figs. 5b, 5c). Note that NFAT is a known interaction partner of FoxP3 and assists FoxP3 binding to DNA only when their binding sites are 3 bp apart, the property used for inferring FoxP3–DNA registers (see Extended Data Fig. 5c for details).

The overall architecture resembled a ladder where the two double-stranded DNA molecules formed side rails bridged by five "rungs", each of which consisted of two FoxP3 subunits bound to different DNA and joined by direct protein-protein interactions ("intra-rung" interactions) (Fig. 2a, 2c). These rungs were separated by 8 bp or 12 bp in an alternating fashion, forming two different types of "inter-rung" interactions (inter-rung^{8bp} and inter-rung^{12bp}) with divergent significance (to be discussed later). Given that both DNA molecules had the helical periodicity of 10.7 bp per turn, this alternating spacing pattern allowed FoxP3 molecules to occupy consecutive major grooves on one side of each DNA. This geometry, in turn, enabled the FoxP3 molecules on opposing DNA to face each other and form the "rungs" of the ladder. None of the intra- and inter-rung interactions resembled the previously reported head-to-head dimerization interaction (Extended Data Fig. 6a)[22], revealing a new mode of molecular assembly for FoxP3.

The two DNA molecules are skew to each other (i.e. non-parallel, non-intersecting). When projected onto the x-y plane as in Fig. 2a, they appeared parallel, but projection onto the x-z plane as in Fig. 2d suggests that they approached each other at an angle of 35˚. The divergence of the two DNA molecules can explain why the multimeric assembly was limited to the decamer spanning ~50 bp near the projected intersection point (Fig. 2d), even though the DNA sample in cryo-EM was 72 bp-long and had many more T_3G repeats to accommodate additional FoxP3 molecules. The lack of cryo-EM density for FoxP3 molecules bound to DNA without forming the "rung" suggests that the intra-rung interaction is critical for stable protein-DNA interaction. In other words, DNA bridging may be an integral part of the assembly.

To test whether DNA bridging indeed occurs in solution, we examined co-purification of non-biotinylated DNA with biotinylated DNA (bait) in the presence and absence of FoxP3. We observed DNA bridging between biotinylated and non-biotinylated T_3G repeats only in the presence of FoxP3$^{\Delta N}$ (Fig. 2e). DNA bridging was not observed between IR-FKHM and IR-FKHM DNAs or between $(T_3G)_{12}$ and IR-FKHM DNAs. Similar T_nG repeat-dependent bridging was observed with full-length FoxP3 expressed in 293T cells (Extended Data Fig. 6b). Additionally, T_3G repeat DNA bridging occurred more efficiently with an increasing concentration of FoxP3 (Extended Data Fig. 6c), suggesting that DNA bridging is not an artificial consequence of saturating multimeric FoxP3 with DNA.

Intra-rung interaction is essential for T_nG repeat recognition, DNA bridging, and cellular functions of FoxP3.

Examination of the intra-rung interaction showed that multiple distinct parts of the protein are involved; wing1 (W1), a loop between helix 2 and 4 (H2/H4 loop) and Helix6 (H6) of one subunit interacted with RBR and H2/H4 of the other subunit within the rung (Fig. 3a). While the resolution at the interface was insufficient to assign precise side chain conformations, the structure identified R356 in the H2/H4 loop, V396 and V398 in W1, D409/E410/F411 in H6 as

residues at the interface (Fig. 3a). We also chose V408 in H6, which was adjacent to the interface residues and is mutated to Met in a subset of IPEX patients[15,28,29]. Mutations of these interface residues, including V408M, disrupted T_3G repeat binding (Fig. 3b) and DNA bridging (Fig. 3c). The same mutations had a minimal impact on IR-FKHM binding (Fig. 3b), which requires head-to-head dimerization of FoxP3[22]. This is consistent with the previous structure showing that these residues are far from either the DNA binding or the head-to-head dimerization interface[22]. The negative effect of the intra-rung mutations on T_3G repeat binding as well as DNA bridging further supports that DNA bridging is required for FoxP3 multimerization on T_3G repeats, rather than a simple consequence of FoxP3 multimerization.

All five intra-rung mutations disrupted cellular functions of FoxP3, as measured by FoxP3-induced gene expression (*e.g. CTLA4* and *CD25*) and T cell-suppressive functions of FoxP3-transduced CD4+ T cells (Fig. 3d, 3e). None of these mutations affected nuclear localization, the level of FoxP3 (Extended Data Figs. 7a, 7b) or FoxP3's interaction with NFAT (Extended Data Fig. 7c), although slight reduction in NFAT binding was seen for V398E. These results suggest that the ladder-like assembly is important for the transcriptional functions of FoxP3.

Architectural flexibility of the ladder-like assembly broadens sequence specificity.

We next examined the potential role of the inter-rung interactions. The inter-rung^{8bp} interaction was mediated by RBR–RBR contacts, which displayed continuous EM density indicative of a strong or ordered interaction (Extended Data Fig. 3f, Fig. 4a). Unlike the intra-rung interface mutations, a mutation in RBR (F331D) disrupted FoxP3 binding to both T_3G repeats and IR-FKHM (Fig. 4b), suggesting that RBR plays an important role in both ladder-like multimerization and head-to-head dimerization [22]. Consistent with the importance of the inter-rung^{8bp} interaction, changes in the inter-rung^{8bp} spacing from 8 bp (1 bp gap) to 9 bp (2 bp gap) or 7 bp (no gap) resulted in a significant impairment in FoxP3 binding to T_3G repeats (Fig. 4c).

In contrast to the inter-rung^{8bp} interaction, cryo-EM density for the inter-rung^{12bp} interaction was difficult to interpret, which could reflect a weak or less-ordered interaction. In keeping with this, FoxP3 binding tolerates a wide range of inter-rung^{12bp} spacings, with equivalent affinity observed for spacings of 11-13 bp (Fig. 4d). Interestingly, while 14-19 bp spacings were not tolerated, DNA with 21-22 bp spacings showed moderate binding. Given that 11-13 bp, 14-19 bp and 21-22 bp spacings would place FoxP3 one, one and a half and two helical turns away from the upstream FoxP3 molecule, this cyclical pattern suggests that the precise positions of FoxP3 are not essential for multimeric assembly, so far as the DNA sequence allows FoxP3 molecules to line up on one side of DNA and form the "rungs". Consistent with the idea, DNA bridging activity showed a similar cyclical pattern (Fig. 4e).

This architectural flexibility may explain our observations in Fig. 1, which showed that FoxP3 could bind to a broad range of T_nG repeat-like sequences besides perfect T_3G repeats. These include tandem repeats of T_2G, T_4G, T_5G and their various mixtures found in the CNR-seq peaks with allelic imbalance (Extended Data Table 1). To examine whether a similar ladder-like architecture forms with T_nG repeat-like sequences that are not perfect T_3G repeats, we used

DNA-bridging activity as a measure of the ladder-like assembly. All 47 pairs of the DNA sequences showing allelic bias in FoxP3 binding *in vivo* and *in vitro* displayed the same allelic bias in DNA bridging (Fig. 4f, Extended Data Table 1). The multimerization-specific IPEX mutation V408 abrogated bridging of T_2G, T_4G, T_5G repeat DNAs (Fig. 4g), suggesting a similar multimeric architecture for FoxP3 regardless of the exact T_nG repeat sequences. Interestingly, suboptimal T_nG repeats (n=2, 4, 5) were bridged with T_3G repeats more efficiently than with themselves (Fig. 4h, Extended Data Fig. 8), suggesting that having a strong DNA as a bridging partner helps FoxP3 binding to suboptimal sequences. These results reveal yet another layer of complexity that can broaden the sequence specificity of FoxP3.

T_nG microsatellite recognition is conserved among FoxP3 orthologs and paralogs.

We next asked the extent of conservation for the ladder-like multimerization on T_nG repeats. We examined FoxP3 from *H. sapiens*, *O. anatinus* and *D. rerio* to compare with *M. musculus* FoxP3 examined above. All three FoxP3 orthologs showed preferential binding to T_3G repeats and IR-FKHM in comparison to a single FKHM or no FKHM (Extended Data Fig. 9a). They also bridged T_3G repeats (Extended Data Fig. 9b), suggesting a ladder-like assembly similar to that of *M. musculus* FoxP3. This is in keeping with the fact that the key residues for multimerization were broadly conserved or interchanged with similar amino acids in FoxP3 orthologs (Extended Data Fig. 9c). Given that *D. rerio* FoxP3 represents one of the most distant orthologs from mammalian FoxP3, these results suggest that T_nG repeat recognition and ladder-like assembly may be ancient properties of FoxP3.

Inspection of sequence alignment of Forkhead TFs revealed that the key residues for multimerization are also well-conserved within the FoxP family, but not outside (Fig. 5a). Biochemical analysis of *M. musculus* FoxP1, FoxP2 and FoxP4 in the FoxP family showed that they preferentially bound T_3G repeats and bridged T_3G repeat DNA as with FoxP3 (Figs. 5b, 5c and Extended Data Fig. 9d). *De novo* motif analysis of previously published ChIP-seq data showed that T_nG repeat-like motif were indeed enriched in FoxP1- and FoxP4-occupied sites (Fig. 5d, the full list and references included in Extended Data Table 4). This feature was particularly strong for FoxP1 in lymphoma cell lines (SU-DHL-6 and U-2932) and mouse neural stem cells—T_nG repeat-like motif was the most significant motif, while FKHM ranked far lower (Fig. 5d). However, in VCap and K-562 cell lines, FoxP1 ChIP-seq peaks did not show T_nG-like elements, although FKHM was identified as one of the most significant motifs in these cells (Extended Data Table 4). Similar context-dependent enrichment of T_nG repeat-like elements was seen with FoxP4 (Fig. 5d, Extended Data Table 4). In contrast, long (>10 nt) T_nG repeat-like elements were not identified from any of the 48 distinct sets of ChIP-seq data for FoxA1, FoxM1, FoxJ2, FoxJ3, FoxQ1 and FoxS1, while FKHM ranked as one of the strongest motifs in many (Extended Data Table 4). These results suggest that preference for T_nG repeat-like sequence and ladder-like assembly are conserved properties of FoxP3 paralogs and orthologs, but may not be shared among all Forkhead TFs.

Discussion

In summary, our findings showed a new mode of TF–DNA interaction that involves TF homo-multimerization and DNA bridging. Upon binding to T_nG repeats, FoxP3 forms a ladder-like multimer, in which FoxP3 utilizes two DNA molecules as scaffolds to facilitate cooperative multimeric assembly. That is, pioneering FoxP3 molecules (possibly a dimer or two dimers with an 8-bp spacing) that bridge DNA would help recruit additional FoxP3 "rungs", which would in turn stabilize the bridged DNA architecture and subsequent rounds of FoxP3 recruitment. Such cooperative assembly enables FoxP3 to preferentially target long repeats of T_nG rather than spurious sequences containing a few copies of T_nG. The DNA bridging activity also implicates FoxP3 as a previously unrecognized class of TF that can directly mediate architectural functions, which may explain the recently observed role of FoxP3 in DNA loop formation[12,13].

How do we reconcile the ladder-like assembly of FoxP3 on T_nG repeats and the previously reported head-to-head dimeric structure on IR-FKHM or related sequences? We speculate that both represent functional forms that may allow FoxP3 to carry out distinct functions depending on DNA sequence. For example, DNA bridging would be a unique consequence of the ladder-like assembly, not shared with the head-to-head dimer. Different structures may also recruit different cofactors based on different surface exposed and differential avidity. This fits the previous microscopic analysis where FoxP3 were found in two distinct types of nuclear clusters associated with different cofactors[16]. Altogether, these findings suggest FoxP3 as a versatile TF that can interpret a wide range of sequences by assembling at least two distinct homo-multimeric structures.

Our findings also implicate functional roles of microsatellites in FoxP TF-mediated transcription regulation. While widely used as genetic tracing markers due to their high degrees of polymorphism, reports of biological functions of microsatellites[30,31], besides its well-known pathogenic roles[32-35], remain sparse[36,37]. Our finding of the T_nG repeat recognition by FoxP3 and other members of the FoxP family, raises the question of whether microsatellites play greater and more direct roles in transcriptional regulation than previously thought. This also prompts speculation that microsatellite polymorphism may contribute to a broad spectrum of diseases through FOXP TF dysregulation, such as autoimmunity via FOXP3, neurodevelopmental disorders via FOXP1, speech and language impairments via FOXP2, and heart and hearing defects via FOXP4.

Data availability

Pull-down seq data have been deposited to the GEO database with the accession code of GSE232754. The structures and cryo-EM maps have been deposited to the PDB and the EMDB under the accession codes of 8SRP and EMD-40737 for the decameric FoxP3 in complex with DNA, and 8SRO and EMD-40736 for the central FoxP3 tetramer in complex with DNA (focused refinement). Other research materials reported here are available on request.

Code availability

The analyses done by Trimmomatic, Bowtie2, Macs2, DiffBind, Bedtools, Meme Fimo, Samtools and Deeptools were on the Harvard Medical School O2 cloud computing platform. The versions of these softwares, as well as their specific parameters and other information, can be found in

supplementary file S1. The conversion of Fimo result to peak bed file is done by Jupyter Notebook. A standalone Jupyter Notebook, with its inputs and expected outputs, as well as a readme file, can be found in supplementary file S1, folder “Fimo_to_bed”. The calculation of peak intensity from raw Deeptools matrix output is done by Jupyter Notebook. A standalone Jupyter Notebook, with its inputs and expected outputs, as well as a readme file, can be found in supplementary file S1, folder “Deeptools_average_calculation”

References

1. Brunkow, M.E., Jeffery, E.W., Hjerrild, K.A., Paeper, B., Clark, L.B., Yasayko, S.A., Wilkinson, J.E., Galas, D., Ziegler, S.F., and Ramsdell, F. (2001). Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. Nat Genet *27*, 68-73. 10.1038/83784.
2. Bennett, C.L., Christie, J., Ramsdell, F., Brunkow, M.E., Ferguson, P.J., Whitesell, L., Kelly, T.E., Saulsbury, F.T., Chance, P.F., and Ochs, H.D. (2001). The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet *27*, 20-21. 10.1038/83713.
3. Fontenot, J.D., Gavin, M.A., and Rudensky, A.Y. (2003). Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. Nat Immunol *4*, 330-336. 10.1038/ni904.
4. Hori, S., Nomura, T., and Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. Science *299*, 1057-1061. 10.1126/science.1079490.
5. Chatila, T.A., Blaeser, F., Ho, N., Lederman, H.M., Voulgaropoulos, C., Helms, C., and Bowcock, A.M. (2000). JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome. J Clin Invest *106*, R75-81. 10.1172/JCI11679.
6. Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X., et al. (2009). Diversity and complexity in DNA recognition by transcription factors. Science *324*, 1720-1723. 10.1126/science.1162327.
7. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650-665. 10.1016/j.cell.2018.01.029.
8. Inukai, S., Kock, K.H., and Bulyk, M.L. (2017). Transcription factor-DNA binding: beyond binding site motifs. Curr Opin Genet Dev *43*, 110-119. 10.1016/j.gde.2017.02.007.
9. Kribelbauer, J.F., Rastogi, C., Bussemaker, H.J., and Mann, R.S. (2019). Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. Annu Rev Cell Dev Biol *35*, 357-379. 10.1146/annurev-cellbio-100617-062719.
10. Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr Opin Genet Dev *43*, 73-81. 10.1016/j.gde.2016.12.007.
11. Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature *527*, 384-388. 10.1038/nature15518.
12. Liu, Z., Lee, D.S., Liang, Y., Zheng, Y., and Dixon, J. (2023). Foxp3 Orchestrates Reorganization of Chromatin Architecture to Establish Regulatory T Cell Identity. bioRxiv. 10.1101/2023.02.22.529589.

13. Ramirez, R.N., Chowdhary, K., Leon, J., Mathis, D., and Benoist, C. (2022). FoxP3 associates with enhancer-promoter loops to regulate T(reg)-specific gene expression. Sci Immunol *7*, eabj9836. 10.1126/sciimmunol.abj9836.
14. van der Veeken, J., Glasner, A., Zhong, Y., Hu, W., Wang, Z.M., Bou-Puerto, R., Charbonnier, L.M., Chatila, T.A., Leslie, C.S., and Rudensky, A.Y. (2020). The Transcription Factor Foxp3 Shapes Regulatory T Cell Identity by Tuning the Activity of trans-Acting Intermediaries. Immunity *53*, 971-984 e975. 10.1016/j.immuni.2020.10.010.
15. Zemmour, D., Charbonnier, L.M., Leon, J., Six, E., Keles, S., Delville, M., Benamar, M., Baris, S., Zuber, J., Chen, K., et al. (2021). Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic CD4(+) T cell perturbations. Nat Immunol *22*, 607-619. 10.1038/s41590-021-00910-8.
16. Kwon, H.-K., Chen, H.-M., Mathis, D., and Benoist, C. (2017). Different molecular complexes that mediate transcriptional induction and repression by FoxP3. Nature Immunology *18*, 1238-1248. 10.1038/ni.3835.
17. Hannenhalli, S., and Kaestner, K.H. (2009). The evolution of Fox genes and their role in development and disease. Nat Rev Genet *10*, 233-240. 10.1038/nrg2523.
18. Benayoun, B.A., Caburet, S., and Veitia, R.A. (2011). Forkhead transcription factors: key players in health and disease. Trends Genet *27*, 224-232. 10.1016/j.tig.2011.03.003.
19. Dai, S., Qu, L., Li, J., and Chen, Y. (2021). Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations. Nucleic Acids Res *49*, 10235-10249. 10.1093/nar/gkab807.
20. Bandukwala, Hozefa S., Wu, Y., Feuerer, M., Chen, Y., Barboza, B., Ghosh, S., Stroud, James C., Benoist, C., Mathis, D., Rao, A., and Chen, L. (2011). Structure of a Domain-Swapped FOXP3 Dimer on DNA and Its Function in Regulatory T Cells. Immunity *34*, 479-491. 10.1016/j.immuni.2011.02.017.
21. Chen, Y., Chen, C., Zhang, Z., Liu, C.-C., Johnson, M.E., Espinoza, C.A., Edsall, L.E., Ren, B., Zhou, X.J., Grant, S.F.A., et al. (2015). DNA binding by FOXP3 domain-swapped dimer suggests mechanisms of long-range chromosomal interactions. Nucleic Acids Research *43*, 1268-1282. 10.1093/nar/gku1373.
22. Leng, F., Zhang, W., Ramirez, R.N., Leon, J., Zhong, Y., Hou, L., Yuki, K., van der Veeken, J., Rudensky, A.Y., Benoist, C., and Hur, S. (2022). The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. Immunity *55*, 1354-1369 e1358. 10.1016/j.immuni.2022.07.002.
23. Kitagawa, Y., Ohkura, N., Kidani, Y., Vandenbon, A., Hirota, K., Kawakami, R., Yasuda, K., Motooka, D., Nakamura, S., Kondo, M., et al. (2017). Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. Nat Immunol *18*, 173-183. 10.1038/ni.3646.
24. Samstein, R.M., Arvey, A., Josefowicz, S.Z., Peng, X., Reynolds, A., Sandstrom, R., Neph, S., Sabo, P., Kim, J.M., Liao, W., et al. (2012). Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. Cell *151*, 153-166. 10.1016/j.cell.2012.06.053.

25. Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327-339. 10.1016/j.cell.2012.12.009.
26. Koh, K.P., Sundrud, M.S., and Rao, A. (2009). Domain requirements and sequence specificity of DNA binding for the forkhead transcription factor FOXP3. PLoS One *4*, e8109. 10.1371/journal.pone.0008109.
27. Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., et al. (2019). The cis-Regulatory Atlas of the Mouse Immune System. Cell *176*, 897-912 e820. 10.1016/j.cell.2018.12.036.
28. Rubio-Cabezas, O., Minton, J.A., Caswell, R., Shield, J.P., Deiss, D., Sumnik, Z., Cayssials, A., Herr, M., Loew, A., Lewis, V., et al. (2009). Clinical heterogeneity in patients with FOXP3 mutations presenting with permanent neonatal diabetes. Diabetes Care *32*, 111-116. 10.2337/dc08-1188.
29. Consonni, F., Ciullini Mannurita, S., and Gambineri, E. (2021). Atypical Presentations of IPEX: Expect the Unexpected. Front Pediatr *9*, 643094. 10.3389/fped.2021.643094.
30. Ibrahim, A., Papin, C., Mohideen-Abdul, K., Le Gras, S., Stoll, I., Bronner, C., Dimitrov, S., Klaholz, B.P., and Hamiche, A. (2021). MeCP2 is a microsatellite binding protein that protects CA repeats from nucleosome invasion. Science *372*. 10.1126/science.abd5581.
31. Contente, A., Dittmer, A., Koch, M.C., Roth, J., and Dobbelstein, M. (2002). A polymorphic microsatellite that mediates induction of PIG3 by p53. Nat Genet *30*, 315-320. 10.1038/ng836.
32. Li, K., Luo, H., Huang, L., Luo, H., and Zhu, X. (2020). Microsatellite instability: a review of what the oncologist should know. Cancer Cell Int *20*, 16. 10.1186/s12935-019-1091-8.
33. Pecina-Slaus, N., Kafka, A., Salamon, I., and Bukovac, A. (2020). Mismatch Repair Pathway, Genome Stability and Cancer. Front Mol Biosci *7*, 122. 10.3389/fmolb.2020.00122.
34. Bonneville, R., Krook, M.A., Kautto, E.A., Miya, J., Wing, M.R., Chen, H.Z., Reeser, J.W., Yu, L., and Roychowdhury, S. (2017). Landscape of Microsatellite Instability Across 39 Cancer Types. JCO Precis Oncol *2017*. 10.1200/PO.17.00073.
35. Kloor, M., and von Knebel Doeberitz, M. (2016). The Immune Biology of Microsatellite-Unstable Cancer. Trends Cancer *2*, 121-133. 10.1016/j.trecan.2016.02.004.
36. Bagshaw, A.T.M. (2017). Functional Mechanisms of Microsatellite DNA in Eukaryotic Genomes. Genome Biol Evol *9*, 2428-2443. 10.1093/gbe/evx164.
37. Gharesouran, J., Hosseinzadeh, H., Ghafouri-Fard, S., Taheri, M., and Rezazadeh, M. (2021). STRs: Ancient Architectures of the Genome beyond the Sequence. J Mol Neurosci *71*, 2441-2455. 10.1007/s12031-021-01850-6.

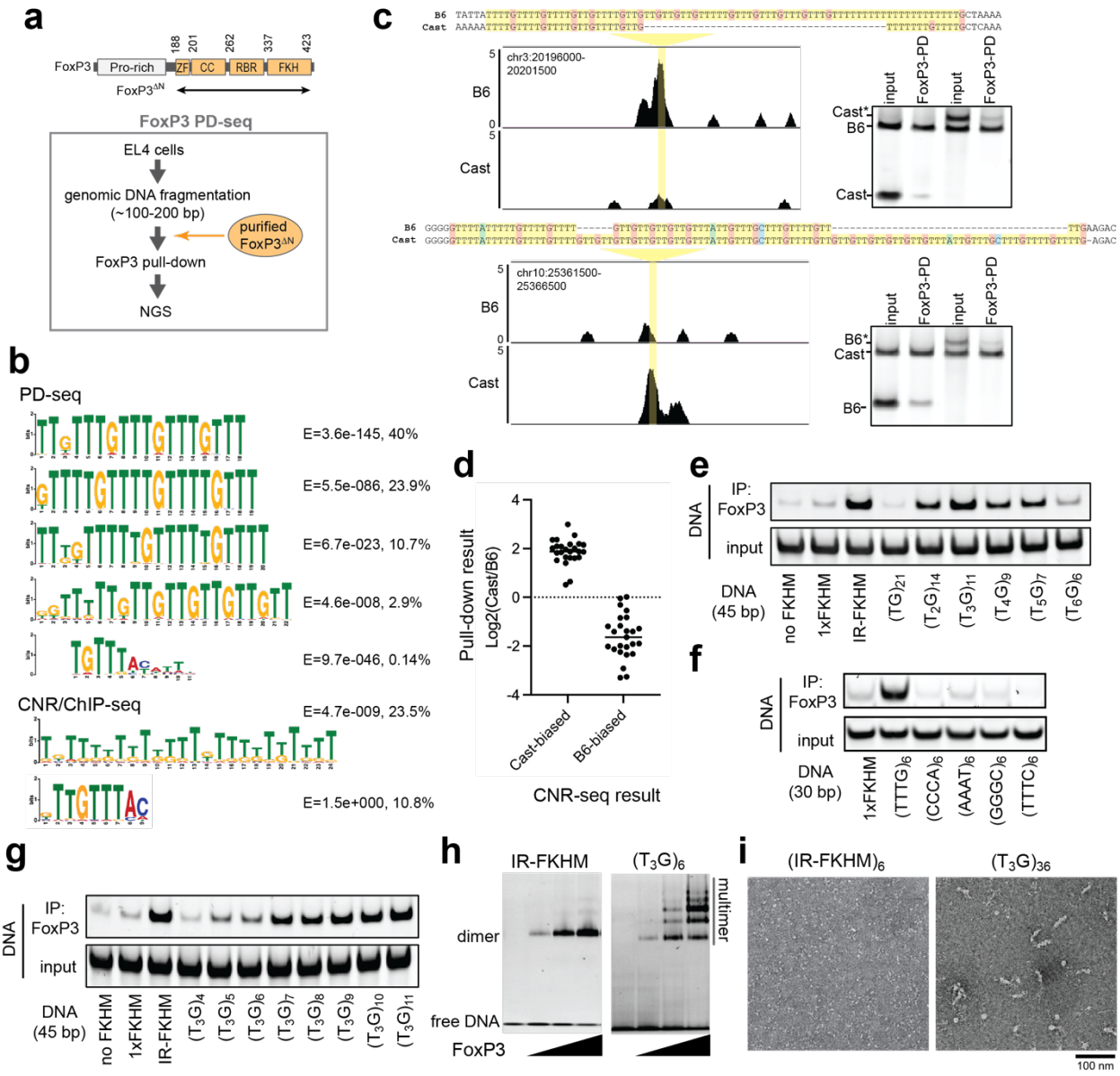


Figure 1. FoxP3 recognizes TnG repeat microsatellites.

a. Domain architecture of FoxP3 and schematic of FoxP3 pull-down (PD)-seq. MBP-tagged FoxP3$^{\Delta N}$ was used for pull-down. ZF: zinc finger; CC: coiled coil; RBR: Runx1-binding region.

b. *De novo* motif analysis of FoxP3 PD-seq peaks (n=21,605) or CNR/ChIP consensus peaks (n=1,754) by MEME-ChIP and STREAM. E-score and percentage of peaks containing the given motif are shown on the right. CNR/ChIP consensus peaks were obtained by identifying overlaps between published FoxP3 CNR-seq and ChIP-seq data (van der Veeken et al., 2020) (see Methods).

c. Allelic imbalance in FoxP3 binding. Left: genome browser view of CNR-seq data (van der Veeken et al., 2020) showing B6-biased (top) and Cast-biased (bottom) peaks. B6 genomic coordinates are shown on the upper left corner. DNA sequences with allelic difference are shown above. Right: FoxP3 pull-down analysis of B6 vs. Cast allele

sequences. B6 and Cast DNA oligos were mixed 1:1, incubated with MBP-FoxP3$^{\Delta N}$ and subjected to MBP pull-down followed by gel analysis. DNA was visualized by SybrGold staining. Cast* and B6* represent Cast and B6 sequences extended with the random sequence (see Extended Data Table 2) to make less-preferred allele sequence longer than preferred allele sequence.

d. FoxP3 affinity comparison between Cast and B6 sequences. 50 pairs of DNA sequences from 50 loci showing allelic bias in FoxP3 CNR-seq (van der Veeken et al., 2020) were chosen (25 each for Cast- and B6-biased peaks, see Extended Data Table 1). Each pair of sequences were subjected to FoxP3$^{\Delta N}$ pull-down as in Fig. 1c. Intensity ratio of Cast to B6 DNA in pull-down eluate was normalized against the input ratio and plotted. Each data point represents an average of the two independent pull-down experiments for a pair of DNA oligos.

e. FoxP3–DNA interaction as measured by FoxP3$^{\Delta N}$ pull-down. DNA harboring a random sequence (no FKHM), a single FKHM (1XFKHM), IR-FKHM or tandem repeats of T_nG (n=1-6) was used. All DNAs were 45 bp-long.

f. FoxP3–DNA interaction using DNAs (30 bp) harboring various tandem repeats, including T_3G repeats.

g. FoxP3–DNA interaction using DNAs (45 bp) harboring 4-11 repeats of T_3G.

h. Native gel shift assay of MBP-FoxP3$^{\Delta N}$ (0, 0.1, 0.2 and 0.4 μM) with DNA (30 bp, 0.05 μM) containing IR-FKHM or $(T_3G)_6$.

i. Representative negative-stain EM images of FoxP3$^{\Delta N}$ in complex with $(T_3G)_{36}$ (top) and $(\text{IR-FKHM})_6$ (bottom). Both DNAs were 144 bp-long. See also Extended Data Fig. 2e.

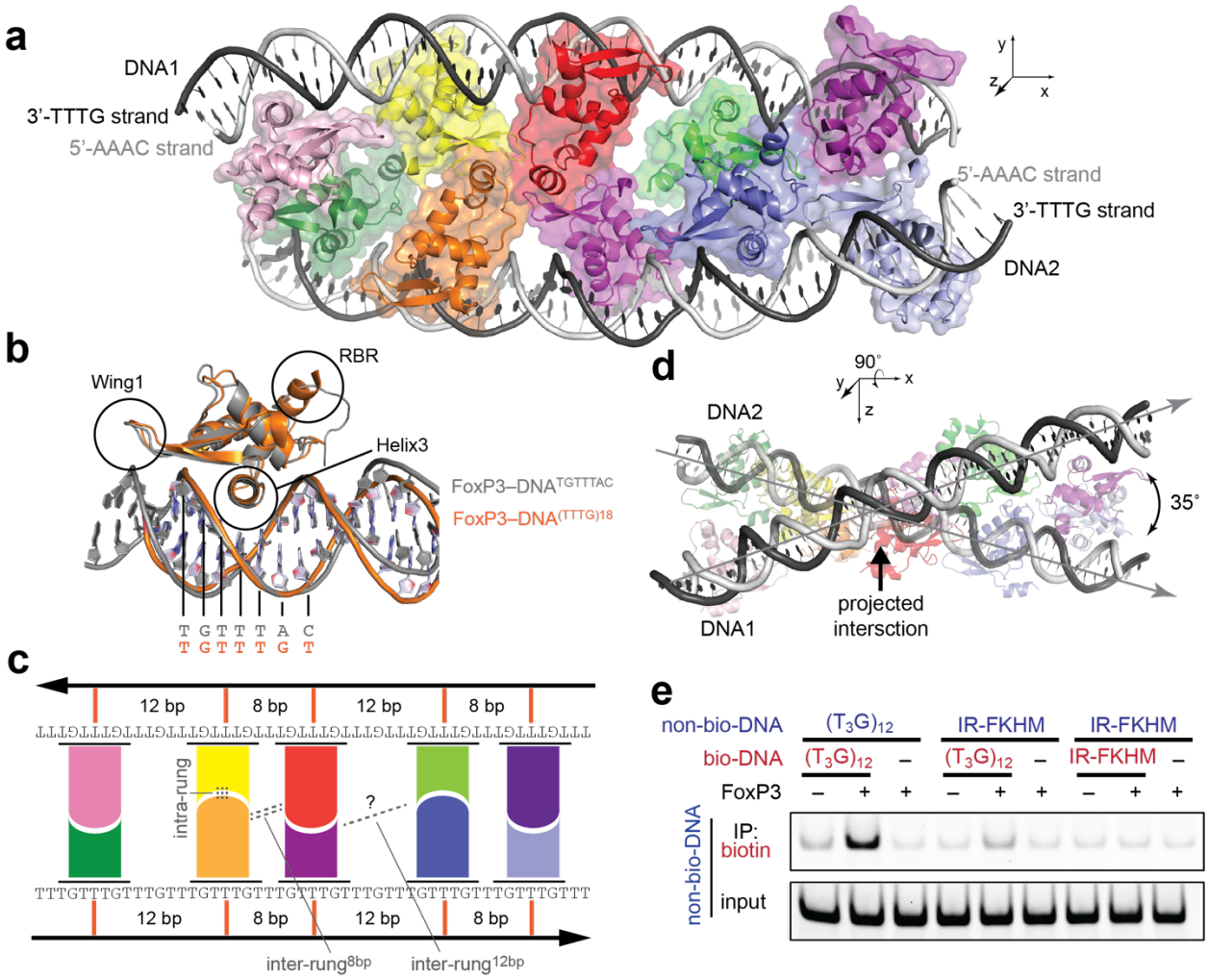


Figure 2. FoxP3 forms a ladder-like multimer upon binding to T_3G repeat DNA.

a. Cryo-EM structure of FoxP3$^{\Delta N}$ decamer in complex with two DNA molecules (grey) harboring $(T_3G)_{18}$. Each of the ten FoxP3 subunits are colored differently.
b. Comparison of a representative FoxP3$^{\Delta N}$ subunit from (a) (orange) with a FoxP3$^{\Delta N}$ subunit from the head-to-head dimeric structure (grey, PDB: 7TDX). Helix3 recognizes DNA sequence (TGTTTAC in the head-to-head dimer, TGTTTGT in the ladder-like multimer) by inserting it into the major groove.
c. Schematic of the ladder-like architecture of FoxP3 on T_3G repeat DNA.
d. Skew relationship between the two DNA molecules, which is evident when looking down the y-axis of (a).
e. DNA bridging assay. Biotinylated DNA (bio-DNA, 82 bp) and non-biotinylated DNA (non-bio-DNA, 60 bp) were mixed at 1:1 ratio (0.1 μM each), incubated with FoxP3$^{\Delta N}$ (0.4 μM), and subjected to streptavidin pull-down prior to gel analysis. Non-biotinylated DNA in the eluate was visualized by SybrGold staining.

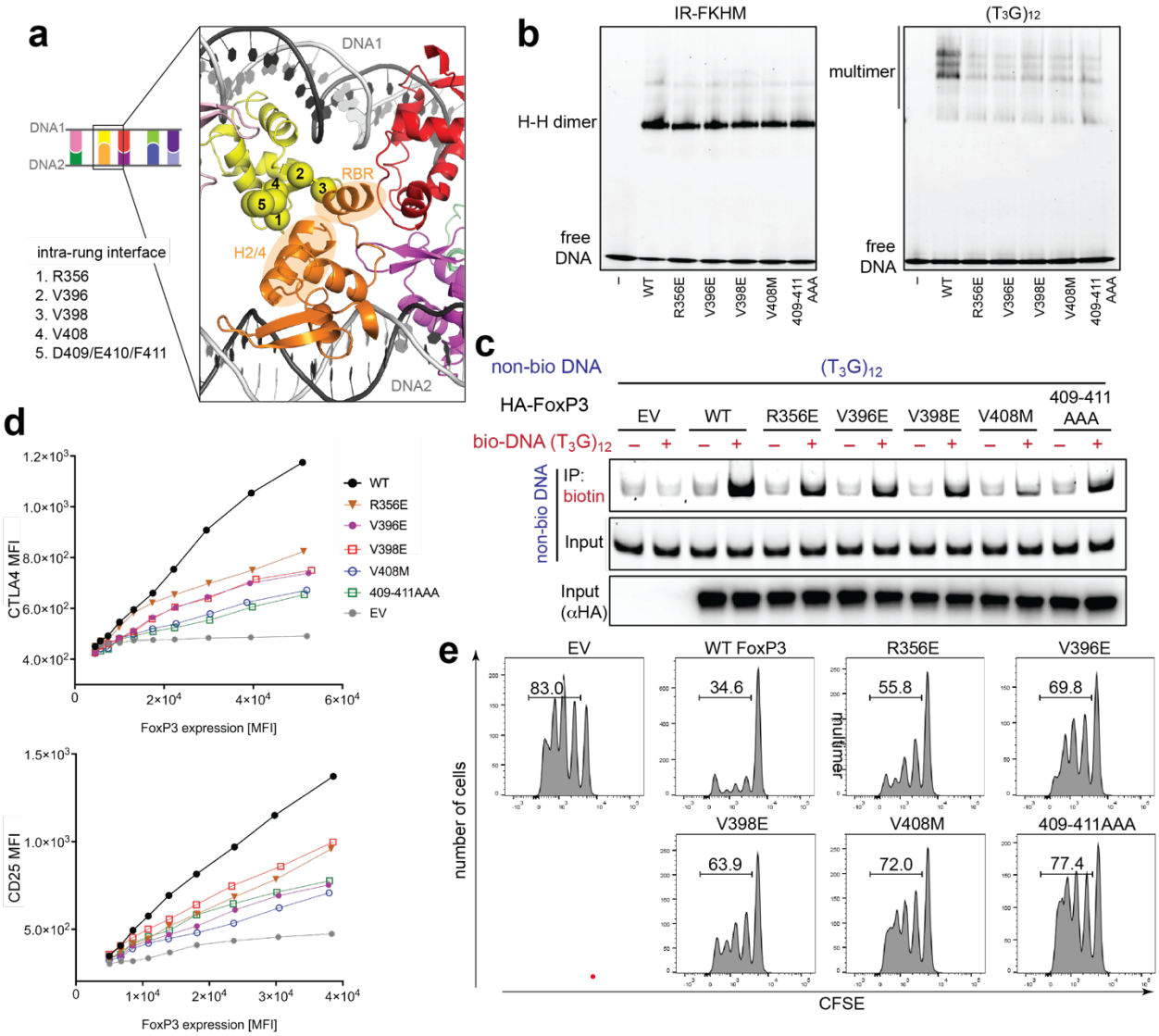


Figure 3. Intra-rung interaction is essential for T_nG repeat recognition, DNA bridging, and cellular functions of FoxP3.

a. Intra-rung interface. Cα of R356, V396, V398, V408 and D409/E410/F411 are shown in spheres. These residues on yellow subunit interact with RBR and helix2-helix4 (H2/4) of the orange subunit. Subunit colors are as in Fig. 2a.

b. Effect of intra-rung interface mutations on DNA binding. MBP-tagged $FoxP3^{\Delta N}$ (0.4 μM) was incubated with IR-FKHM or $(T_3G)_{12}$ (60 bp for both) and analyzed by native gel shift assay.

c. Effect of intra-rung interface mutations on DNA bridging. FoxP3 (or empty vector, EV) was expressed in 293T cells and lysate was incubated with a mixture of biotinylated and non-biotinylated DNA (1:1 ratio), followed by streptavidin pull-down and gel analysis.

d. Transcriptional activity of FoxP3. CD4+ T cells were retrovirally transduced to express FoxP3, and its transcriptional activity was analyzed by measuring the level of the known target genes CTLA4 and CD25 using FACS. FoxP3 level was measured using Thy 1.1, which is under the control of IRES from the bicistronic mRNA for FoxP3.

e. T cell suppression assay. FoxP3-transduced T cells (suppressors) were mixed with naïve T cells (responders) at a 1:2 ratio and the effect of the suppressor cells on proliferation of the responder cells were measured by the CFSE dilution profile of the responder T cells.

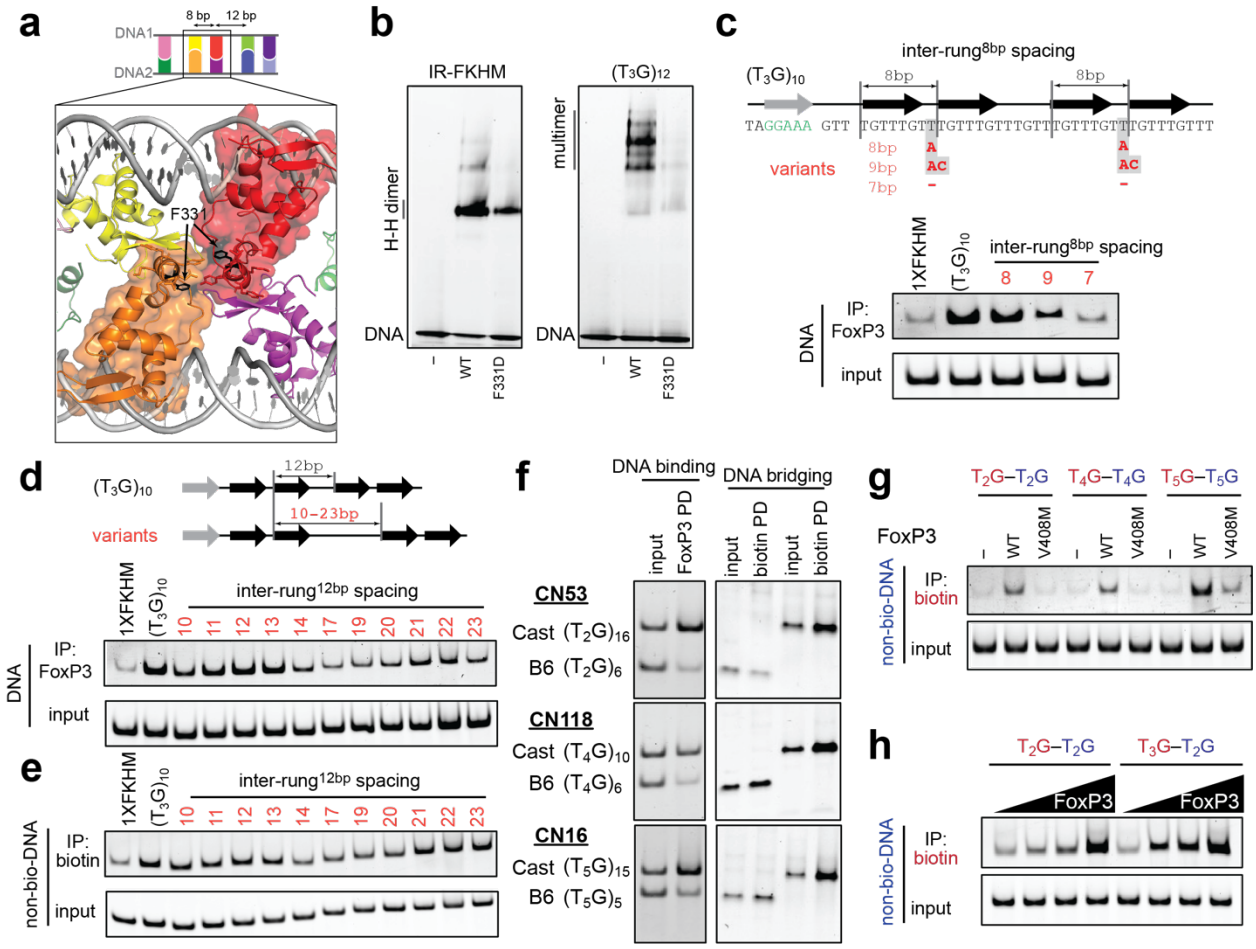


Figure 4. Architectural flexibility of the ladder-like assembly broadens sequence specificity of FoxP3.

a. Structure highlighting the inter-rung^{8bp} interaction between orange and red subunits (in surface representation). The interaction is primarily between RBRs, where F331 constitutes one of the interface residues.

b. Effect of inter-rung^{8bp} mutation (F331D) on FoxP3–DNA interaction. Experiment was done as in Fig. 3b.

c. Effect of inter-rung^{8bp} spacing variations on FoxP3 binding as measured by FoxP3 pull-down. Both inter-rung^{8bp} gap nucleotides were changed from T in $(T_3G)_{10}$ to A (without changing the 8-bp spacing), to AC (9-bp spacing) or to no nucleotide (7-bp spacing). Black arrows indicate FoxP3 footprints. Gray arrow indicates the NFAT binding site. NFAT interacts with FoxP3 and helps fixing the FoxP3-DNA register to one particular configuration, which was necessary to examine the effect of DNA sequence variations at or between the FoxP3 footprints.

d. Effect of inter-rung^{12bp} spacing variations on FoxP3 binding. The inter-rung^{12bp} spacing was changed from 12 bp in $(T_3G)_{10}$ to 10-23 bp. See Extended Data Table 2 for sequences.

e. Effect of inter-rung^{12bp} spacing variations on FoxP3-mediated DNA bridging. Non-biotinylated DNAs in (d) were mixed with biotinylated $(T_3G)_{10}$ and FoxP3$^{\Delta N}$ (0.2 μM) prior to streptavidin pull-down and gel analysis.

f. Comparison of Cast and B6 sequences in FoxP3 binding (left) and DNA bridging (right). Three pairs of sequences at loci CN53, CN118 and CN16 that showed Cast-bias in CNR-seq were compared (see Extended Data Table S1 for the sequences and compilation of the results for the entire 50 loci). Experiments were performed as in (d) and (e).
g. DNA bridging between $(T_2G)_{14}$ and $(T_2G)_{14}$, between $(T_3G)_{11}$ and $(T_3G)_{11}$, and between $(T_4G)_9$ and $(T_4G)_9$ in the presence of WT FoxP3 or the IPEX mutant V408M. Biotinylated and non-biotinylated DNA were colored red and blue, respectively.
h. DNA bridging between $(T_2G)_{14}$ and $(T_2G)_{14}$, and between $(T_2G)_{14}$ and $(T_3G)_{11}$ with an increasing concentration of FoxP3 (0, 0.1, 0.2, 0.4 μM).

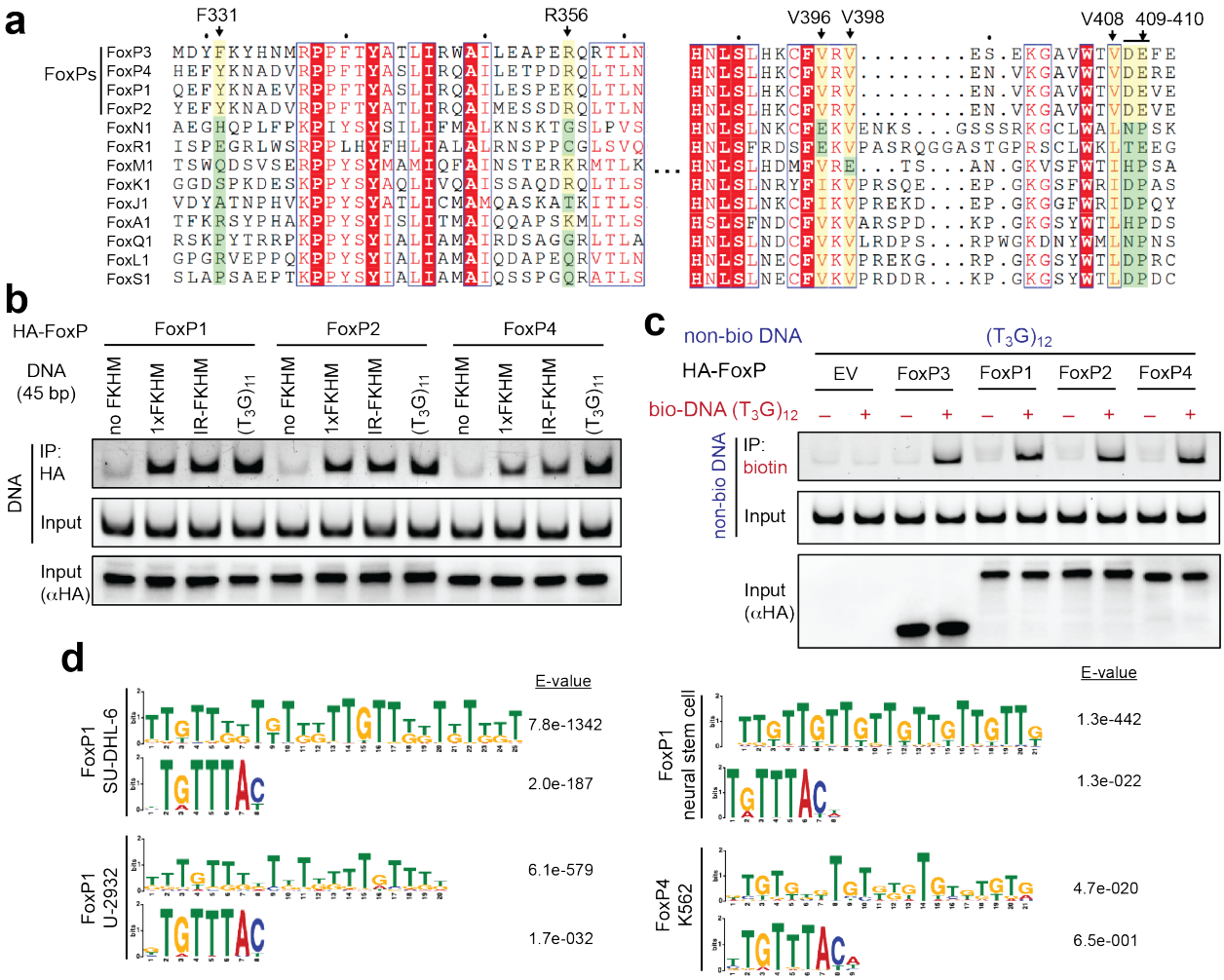


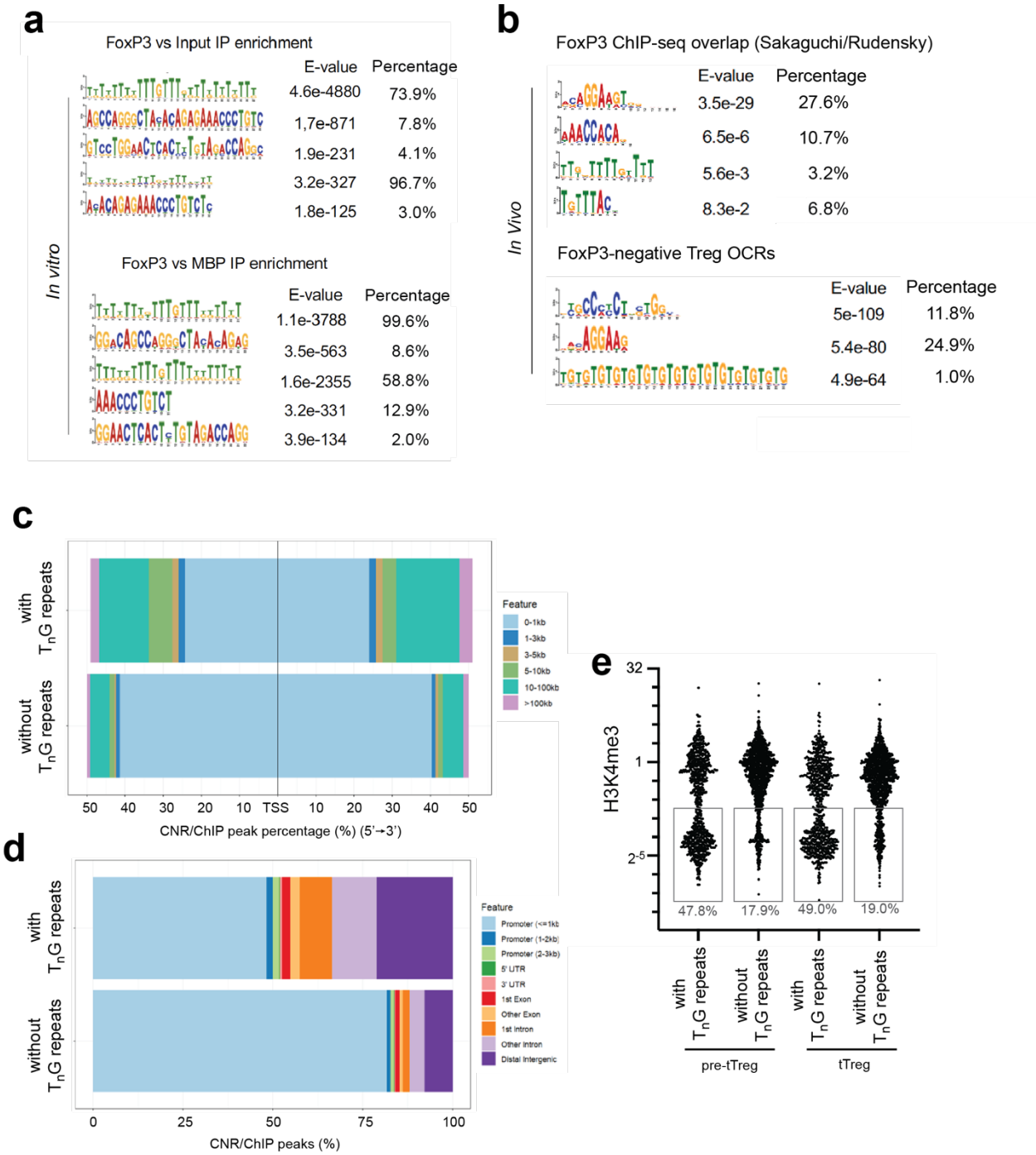
Figure 5. T_nG microsatellite recognition is conserved among FoxP3 orthologs and paralogs.

a. Sequence alignment of forkhead TFs. Residues equivalent to the key interface residues in mouse FoxP3 (arrow on top) were highlighted yellow (when similar to the mouse residues) or green (when dissimilar).

b. DNA-binding activity of FoxP1, FoxP2 and FoxP4. HA-tagged FoxP1, FoxP2 and FoxP4 were transiently expressed in HEK293T cells and purified by anti-HA IP. Equivalent amounts of indicated DNAs (all 45 bp) were added to FoxP1/2/4-bound beads and further purified prior to gel analysis (SybrGold).

c. DNA-bridging activity of FoxP1, FoxP2 and FoxP4. Experiments were performed as in Fig. 3c using 293T lysate expressing HA-tagged FoxP TFs.

d. *De novo* motif analysis of FoxP1 and FoxP4 ChIP-seq peaks from published database (Braccioli et al., 2017; Consortium, 2012). See Extended Data Table 4 for the comprehensive list and their references.



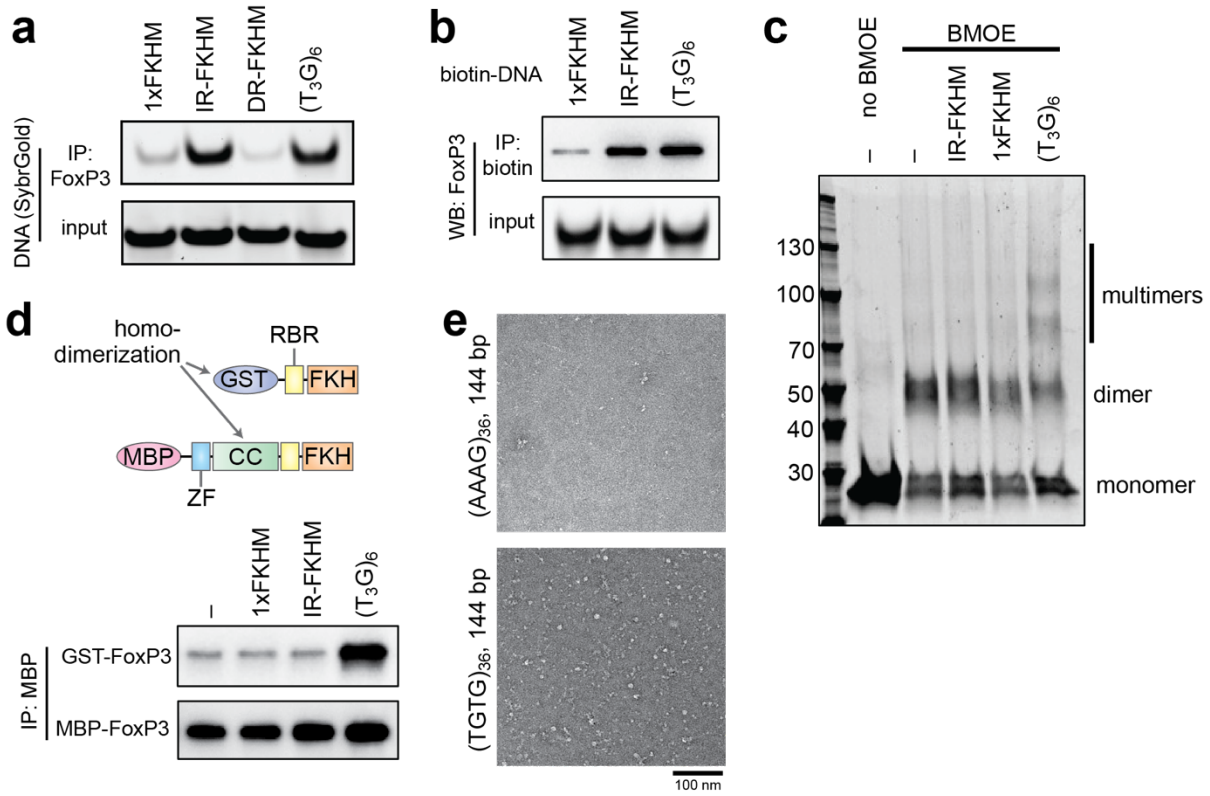
Extended Data Figure 1. Analysis of FoxP3 PD-seq and CNR/ChIP-seq peaks

a. *De novo* motif analysis of MBP-FoxP3 PD-seq peaks vs. Input (n = 21,605) or MBP-FoxP3 PD-seq vs. MBP alone PD-seq peaks (n = 13,039) using MEME-ChIP. E-value and percentage of peaks harboring the given motif are shown on the right.

b. Top: *De novo* motif analysis of FoxP3 ChIP-seq overlapping peaks from (Kitagawa et al., 2017) and (Samstein et al., 2012) (n = 5,000). Bottom: *de novo* motifs in FoxP3-negative Treg OCRs (n = 216,419), which were defined from all observed Treg ATAC-seq peaks

(Yoshida et al., 2019) depleted of FoxP3 ChIP-seq peaks. Overlap peaks and depleted peaks were calculated using bedtools (see Methods).

c. Distribution of CNR/ChIP peaks (as defined in Fig. 1b) relative to Transcription Start Sites (TSS). CNR/ChIP peaks harboring sequences that match the T_nG repeat-like motif in Fig. 1b (n = 716) or peaks without T_nG repeat-like motif (n = 1,038) were compared. FIMO (p = 0.05) was used to identify peaks with and without T_nG repeat-like sequences.
d. Distribution of the genomic features in CNR/ChIP peaks with (n = 716) and without T_nG repeat-like motif (n= 1,038).
e. Level of H3K4me3-ChIP signal (Kitagawa et al., 2017) around the CNR/ChIP peaks with (n = 716) and without T_nG repeat-like motif (n= 1,038). H3K4me3 ChIP signal was averaged over +/- 2 kb around the CNR peak summits.



Extended Data Figure 2. FoxP3 forms multimers on T_3G repeats.

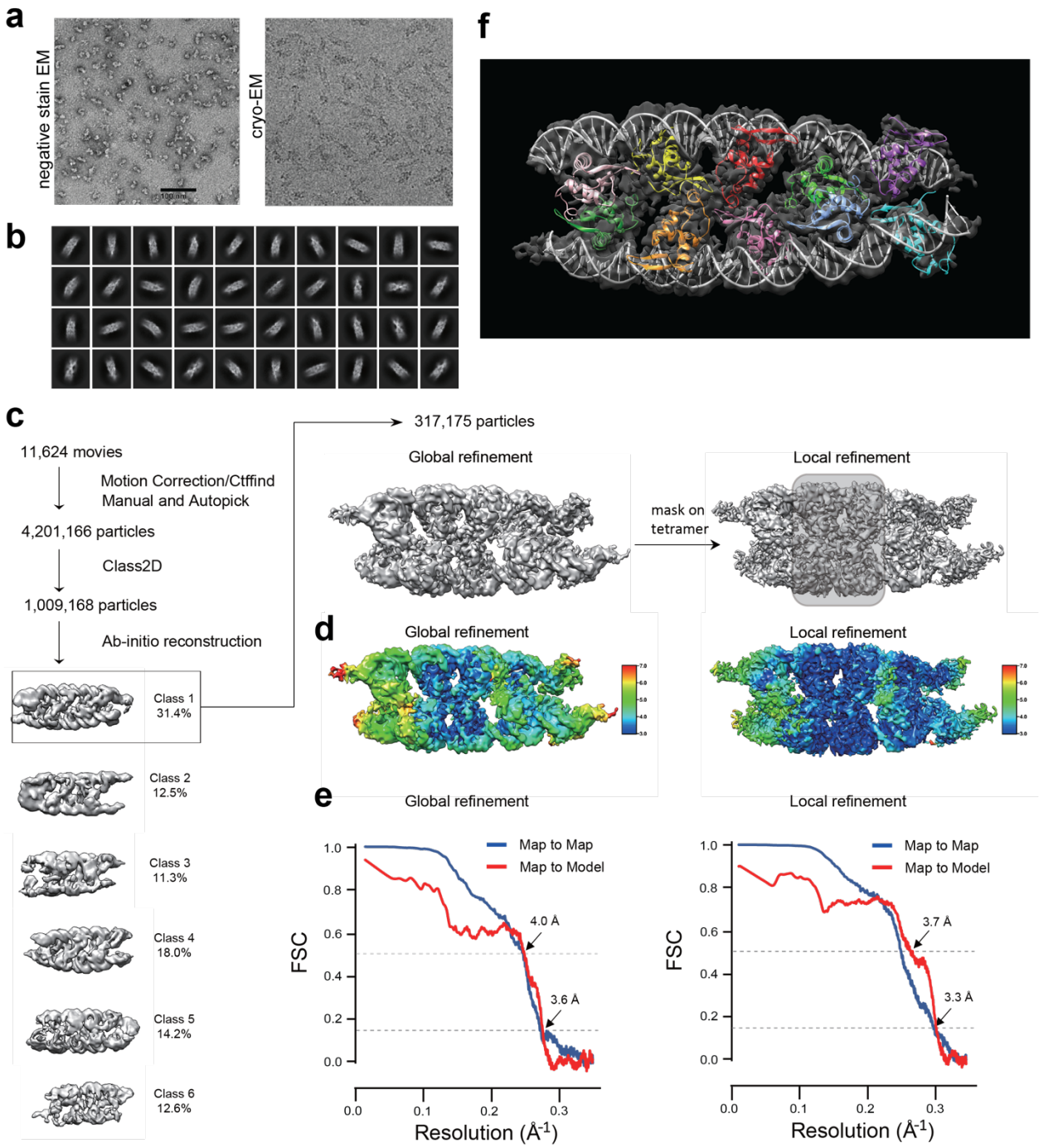
a. DNA sequence specificity of FoxP3 as measured by FoxP3 pull-down. HA-tagged, full-length FoxP3 was transiently expressed in HEK293T cells and purified by anti-HA IP. Equivalent amounts of indicated DNAs (30-31 bp) were added to FoxP3-bound beads and further purified by anti-HA IP prior to gel analysis.

b. DNA sequence specificity of FoxP3 as measured by DNA pull-down. Equivalent amounts of biotinylated DNAs were mixed with FoxP3-expressing 293T lysate and were subjected to streptavidin pull-down. Co-purified FoxP3 was analyzed by anti-HA WB.

c. BMOE crosslinking of $FoxP3^{\Delta N}$ with and without DNA. $FoxP3^{\Delta N}$ can only form multimers on $(T_3G)_6$ DNA.

d. Multimerization analysis of FoxP3, as measured by co-purification of FoxP3 with different tags. GST- and MBP-tagged FoxP3 were incubated together in the presence and absence of indicated DNA and were subjected to MBP pull-down, followed by WB analysis of GST-FoxP3 in eluate. Note that GST replaced the CC domain in FoxP3, disallowing hetero-dimerization between MBP-FoxP3 and GST-FoxP3. Thus, co-purification of these two proteins in the presence of T_3G repeats suggests DNA sequence-dependent multimerization of the FoxP3 homodimer.

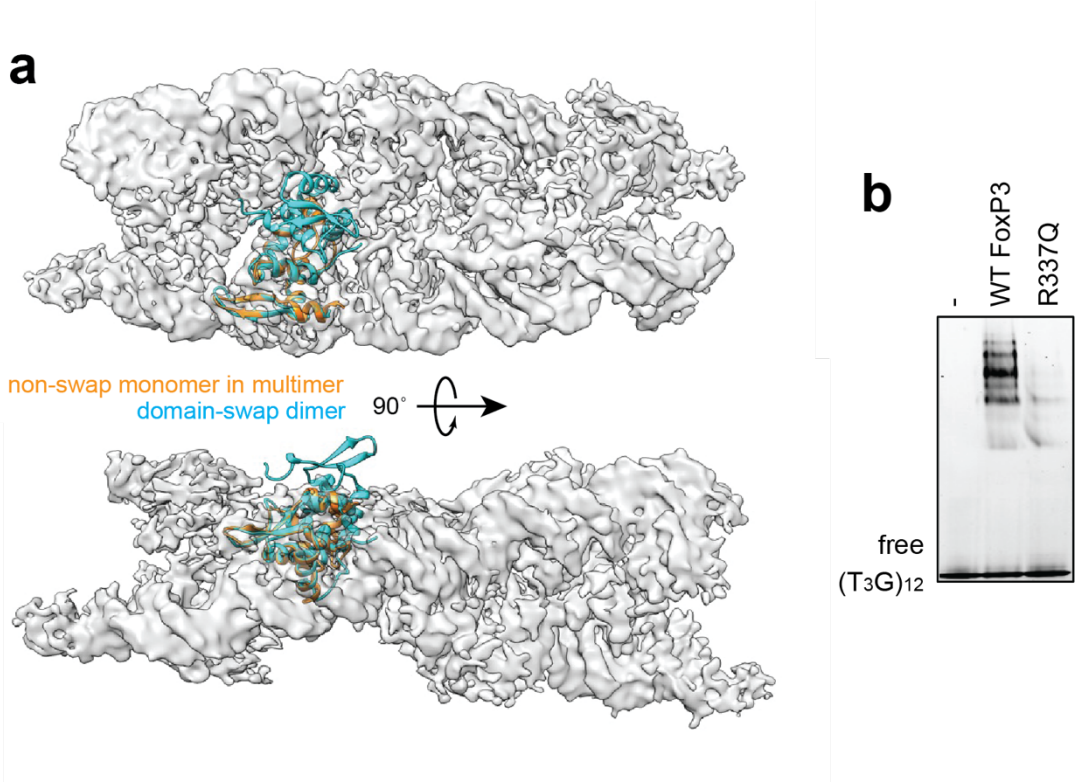
e. Representative negative-stain EM images of $FoxP3^{\Delta N}$ in complex with $(AAAG)_{36}$ (top) and $(TGTG)_{36}$ (bottom).



Extended Data Figure 3. Cryo-EM structure determination of the FoxP3$^{\Delta N}$–$(T_3G)_{18}$ complex.

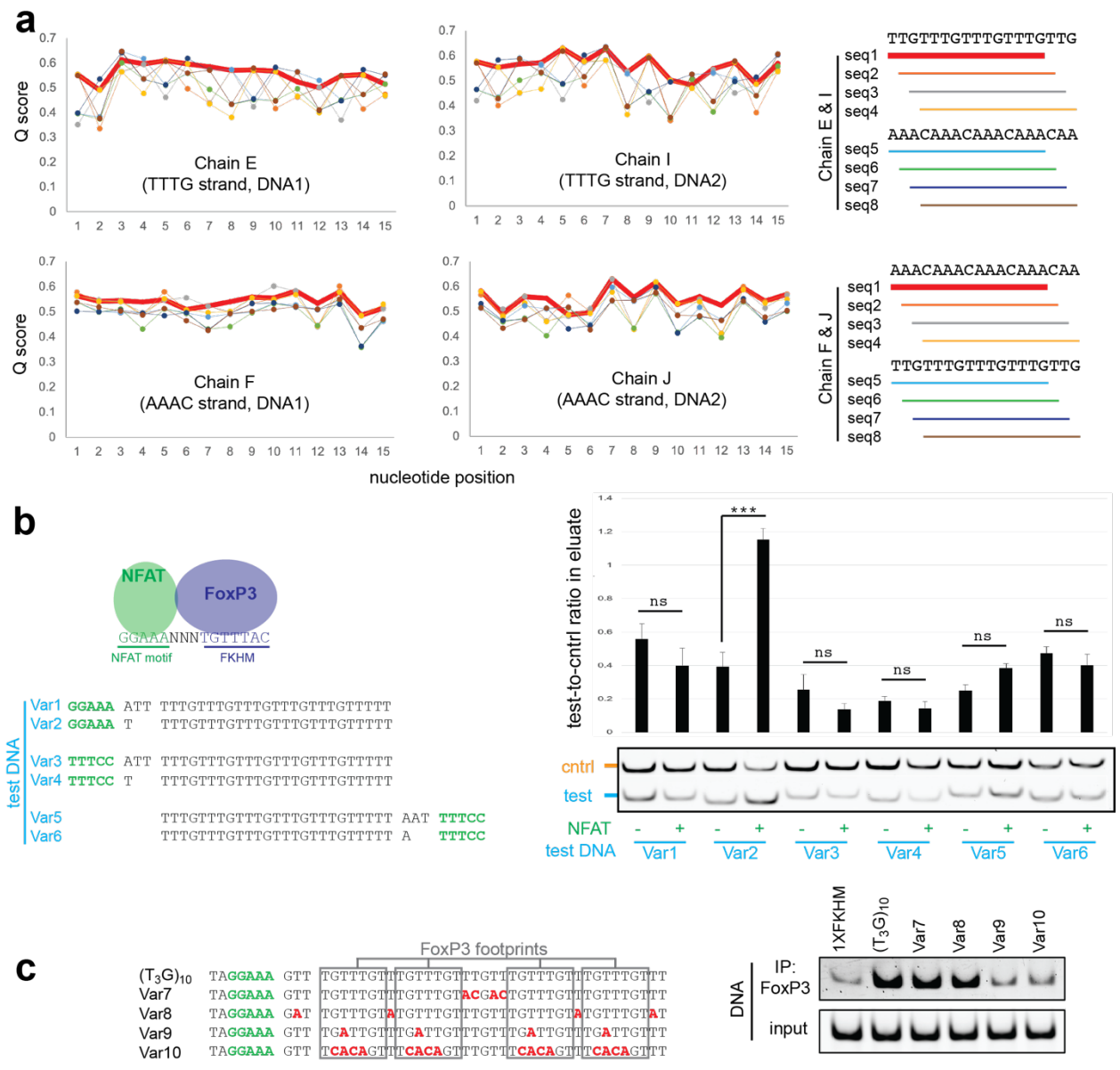
a. Representative negative-stain EM (left) and cryo-EM images (right) of FoxP3$^{\Delta N}$ multimers on $(T_3G)_{18}$ DNA.
b. 2D classes chosen for 3D reconstruction.
c. Cryo-EM image processing workflow. See details in Methods.

d. Local resolution for the maps of global refinement (left) and local refinement with a mask covering the central four subunits of FoxP3 (right). Local resolution was calculated by CryoSPARC. Resolution range was indicated according to the color bar.
e. Fourier shell correlation (FSC) curve for global refinement (left) and local refinement (right). Map-to-Map FSC curve was calculated between the two independently refined half-maps after masking (blue line), and the overall resolution was determined by gold standard FSC=0.143 criterion. Map-to-Model FSC was calculated between the refined atomic models and maps (red line).
f. Cryo-EM map and ribbon model of $FoxP3^{\Delta N}$ decamer in complex with two $(T_3G)_{18}$ DNAs (PDB: 8SRP, EMDB: 40737). DNA molecules are colored grey. Individual FoxP3 monomers are colored differently.



Extended Data Figure 4. Domain-swap dimerization is incompatible with FoxP3 binding to T_3G repeats.

a. Superposition of the domain-swap dimeric structure of FoxP3 (cyan, PDB:4WK8) onto any subunit of the FoxP3 multimeric structure (represented here by the orange subunit) by aligning the common portions of FoxP3 reveals that the domain-swap dimer is incompatible with the density map.
b. Native gel shift analysis of MBP-tagged $FoxP3^{\Delta N}$ (WT or R337Q, 0.4 μM) with $(T_3G)_{12}$ DNA (0.05 μM). Note that R337Q induces domain-swap dimerization.



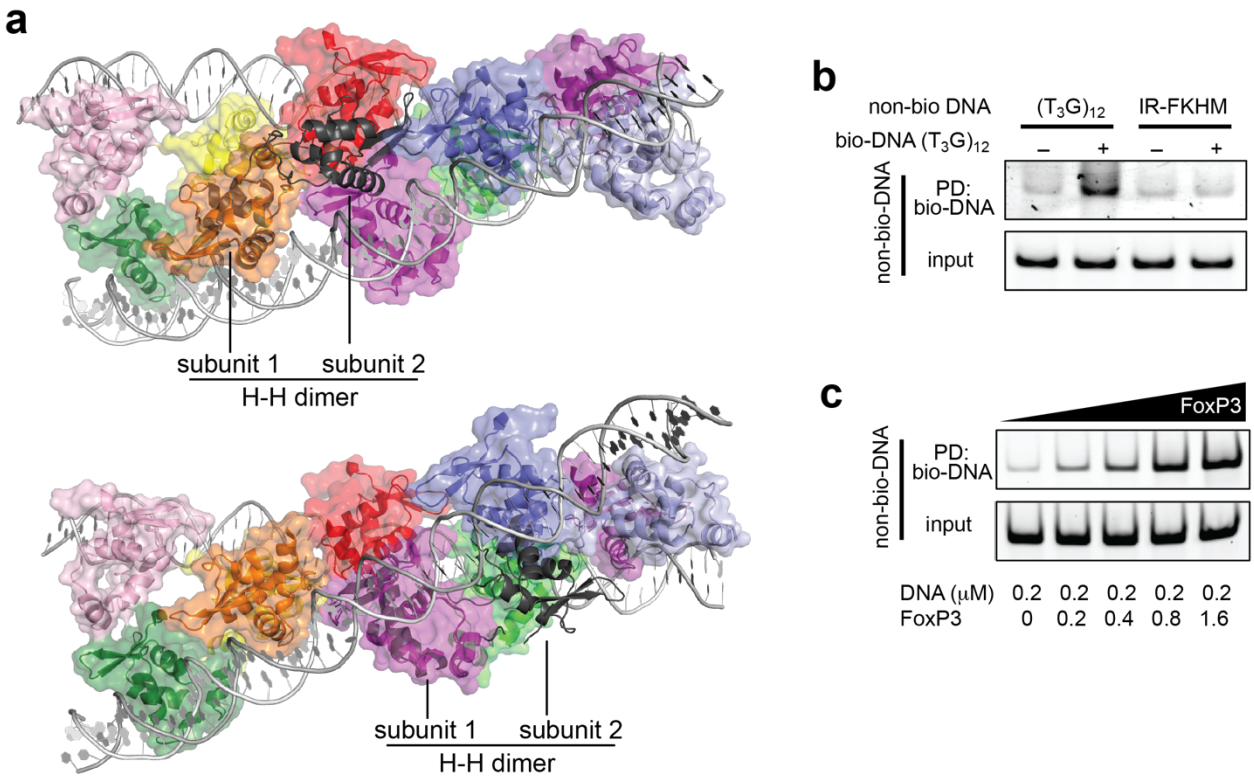
Extended Data Figure 5. Sequence assignment of DNA in the cryo-EM structure.

a. DNA sequence assignment by Q-score analysis using mapQ. For each of the four DNA strands, eight possible sequence alignments (right) were tested against the local refinement map. The sequence alignment showing the highest overall score was highlighted with a thick red line. The best sequence alignments for all four DNA strands were consistent with each other and were used for cryo-EM reconstruction.

b. Experimental validation of FoxP3–DNA registers using NFAT–FoxP3 cooperativity analysis. This assay utilizes the FoxP3 interaction partner NFAT, which assists FoxP3 binding to DNA only when their binding sites are 3 bp apart in one particular orientation (as in the schematic). To investigate FoxP3 footprints on T_3G repeat DNA, we varied the position and orientation of NFAT consensus sequence (GGAAA, green) relative to T_3G repeats (Var1-6), and performed FoxP3 pull-down. An internal control DNA (cntrl,

harboring the NFAT motif followed by FKHM with 3 nt gap) was used to normalize the test DNA (Var1-6) pull-down efficiency. Only DNA with a single nucleotide gap between GGAAA and TTTG (Var2) showed a positive effect of NFAT on the FoxP3–DNA interaction. This suggests that the most upstream FoxP3 subunit recognizes TGTTTGT.

c. FoxP3 interaction with $(T_3G)_{10}$ variants. Variations in DNA sequence outside the FoxP3 footprints were tolerated (Var7 and Var8), but those within the footprints (Var9 and Var10) were not.

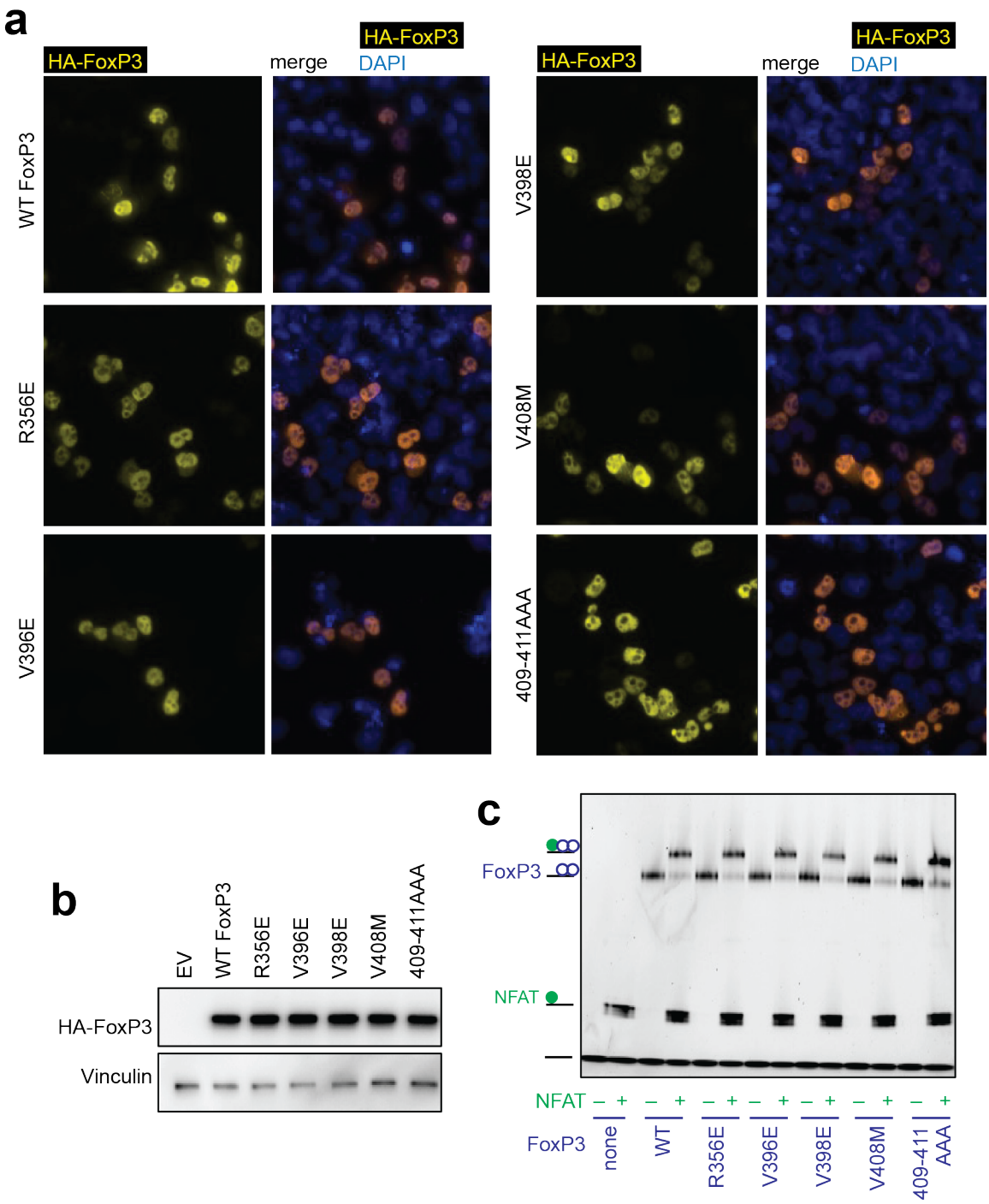


Extended Data Figure 6. Analysis of the inter-subunit interactions and DNA bridging

a. Comparison of the inter-subunit interactions in FoxP3 decamer on T_nG repeats vs. head-to-head (H-H) dimer on IR-FKHM (PDB:7TDX). Superposition of the H-H dimer (dark grey) onto any of the ten subunits in the decamer structure showed distinct modes of inter-subunit interactions. Shown are two examples where subunit 1 of the H-H dimer was aligned to orange (top) or magenta (bottom) subunits of the decamer, showing that subunit 2 of the H-H dimer did not align with any of the decamer subunits.

b. DNA bridging assay using FoxP3 expressed in 293T cells. Biotinylated and non-biotinylated DNA (82 and 60 bp, respectively) were mixed at 1:1 ratio and further incubated with 293T lysates, followed by streptavidin pull-down and gel analysis of non-biotinylated DNA by SybrGold staining.

c. DNA bridging assay using an increasing concentration of purified $FoxP3^{\Delta N}$. 0.1 μM each of biotinylated and non-biotinylated $(T_3G)_{12}$ DNAs were used. FoxP3 concentrations are indicated at the bottom.



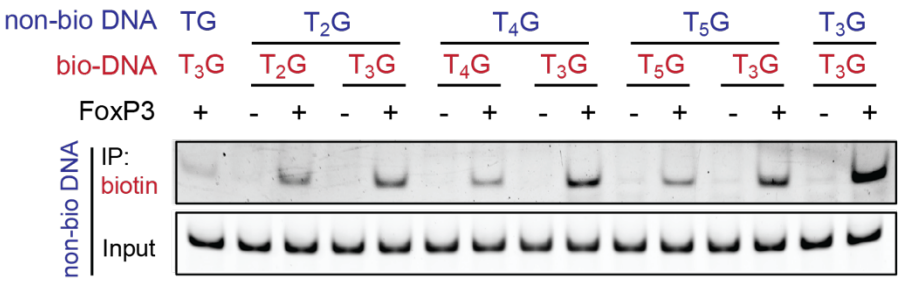
Extended Data Figure 7. Characterization of intra-rung interface mutant FoxP3

a. Nuclear localization of WT FoxP3 and intra-rung interface mutants. HA-tagged FoxP3 was transiently expressed in A549 cells and was subjected to anti-HA immunofluorescent (yellow) analysis. Nuclei were shown with DAPI (blue) staining.

b. Expression levels of WT FoxP3 and intra-rung mutants in A549 cells.

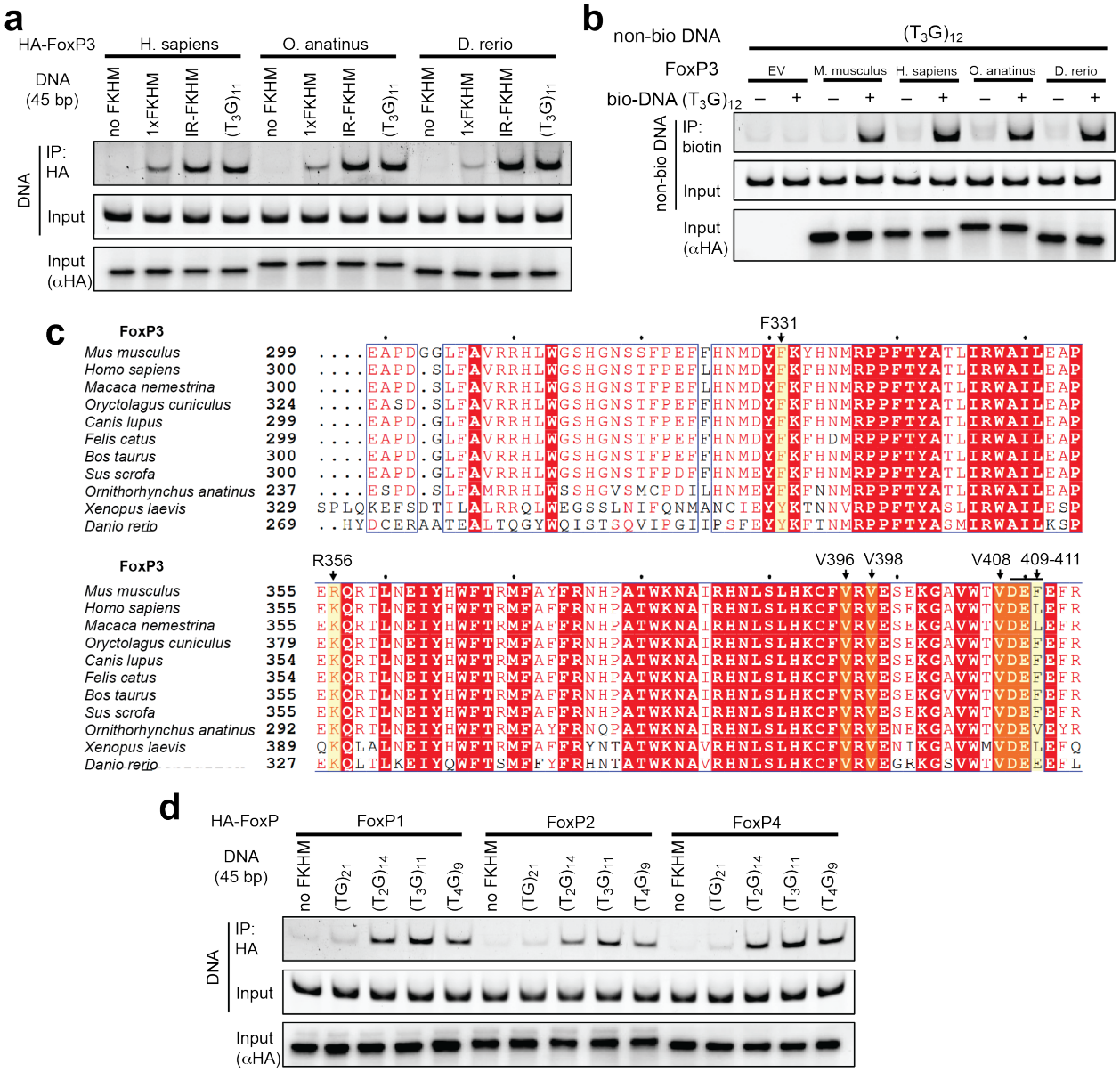
c. Effect of the intra-rung mutations on the NFAT–FoxP3 interaction, as measured by native gel shift assay. FoxP3 (0.1 μM) was incubated with DNA harboring IR-FKHM

and the NFAT site (with a 3-bp gap as in Extended Data Fig. 5b, 0.05 μM). NFAT (0.1 μM) was added to the mixture to monitor formation of the ternary complex NFAT–FoxP3–DNA.



Extended Data Figure 8. Comparison between homo- and hetero-typic DNA-bridging by FoxP3

DNA-bridging activity of FoxP3 with different combinations of T_nG repeats. Biotinylated-DNA (red) and non-biotinylated DNA (blue) were mixed at 1:1 ratio and were incubated with $FoxP3^{\Delta N}$ (0.2 μM) prior to streptavidin pull-down and gel analysis of non-biotinylated DNA. T_2G, T_4G and T_5G repeats bridged better with T_3G repeats than with themselves.



Extended Data Figure 9. Multimerization on T_nG repeats is conserved in FoxP3 orthologs and paralogs.

- **a.** DNA-binding activity of FoxP3 orthologs with indicated DNA. Experiments were performed as in Fig. 5b.
- **b.** DNA-bridging activity of FoxP3 orthologs. Experiments were performed as in Fig. 5c.
- **c.** Sequence alignment of FoxP3 orthologs from different species, showing conservation of the key interface residues (yellow highlight, arrows on top with the residue identities in *M. musculus* FoxP3).
- **d.** DNA-bridging activity of FoxP3 paralogs with T_nG repeats (n=1-4). Experiments were performed as in Fig. 5c.

References

Braccioli, L., Vervoort, S.J., Adolfs, Y., Heijnen, C.J., Basak, O., Pasterkamp, R.J., Nijboer, C.H., Coffer, P.J., 2017. FOXP1 Promotes Embryonic Neural Stem Cell Differentiation by Repressing Jagged1 Expression. Stem Cell Reports 9, 1530-1545.

Consortium, E.P., 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.

Kitagawa, Y., Ohkura, N., Kidani, Y., Vandenbon, A., Hirota, K., Kawakami, R., Yasuda, K., Motooka, D., Nakamura, S., Kondo, M., Taniuchi, I., Kohwi-Shigematsu, T., Sakaguchi, S., 2017. Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. Nat Immunol 18, 173-183.

Samstein, R.M., Arvey, A., Josefowicz, S.Z., Peng, X., Reynolds, A., Sandstrom, R., Neph, S., Sabo, P., Kim, J.M., Liao, W., Li, M.O., Leslie, C., Stamatoyannopoulos, J.A., Rudensky, A.Y., 2012. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. Cell 151, 153-166.

van der Veeken, J., Glasner, A., Zhong, Y., Hu, W., Wang, Z.M., Bou-Puerto, R., Charbonnier, L.M., Chatila, T.A., Leslie, C.S., Rudensky, A.Y., 2020. The Transcription Factor Foxp3 Shapes Regulatory T Cell Identity by Tuning the Activity of trans-Acting Intermediaries. Immunity 53, 971-984 e975.

Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., Tellier, J., Kim, E., Dwyer, D., Shinton, S., Nabekura, T., Qi, Y., Yu, B., Robinette, M., Kim, K.W., Wagers, A., Rhoads, A., Nutt, S.L., Brown, B.D., Mostafavi, S., Buenrostro, J.D., Benoist, C., Immunological Genome, P., 2019. The cis-Regulatory Atlas of the Mouse Immune System. Cell 176, 897-912 e820.