The impact of repetitive DNA and its guardian proteins on the evolution of neuronal gene regulatory networks

Farmiloe, G.

Publication date
2023

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns.

Grace Farmiloe*¹, Gerrald A. Lodewijk*¹, Stijn F. Robben*¹, Elisabeth J. van Bree¹, Frank M.J. Jacobs¹#

¹University of Amsterdam, Swammerdam Institute for Life Sciences, Amsterdam Neuroscience, Amsterdam, the Netherlands
*These authors contributed equally
#Corresponding author

Published: 10 February 2020
Philosophical transactions of the royal society B
https://doi.org/10.1098/rstb.2019.0333
Abstract

The large family of KRAB zinc finger (KZNF) genes are transcription factors (TFs) implicated in recognizing and repressing repetitive sequences such as transposable elements (TEs) in our genome. Through successive waves of retrotransposition mediated insertions, various classes of TEs have invaded mammalian genomes at multiple timepoints throughout evolution. Even though most of the TE-classes in our genome lost the capability to retrotranspose millions of years ago, it remains elusive why the KZNFs that evolved to repress them are still retained in our genome. One hypothesis is that KZNFs become repurposed for other regulatory roles. Here, we find evidence that evolutionary changes in KZNFs provide them not only with the ability to repress TEs, but also to bind to gene promoters independent of TEs. Using KZNF binding site data in conjunction with gene expression values from the Allen Brain Atlas we show that KZNFs have the ability to regulate gene expression in the human brain in a region-specific manner. Our analysis shows that the expression of KZNFs show correlations with the expression of their target genes, suggesting that KZNFs have a direct influence on gene expression in the developing human brain. The extent of this regulation and the impact it has on primate brain evolution is still to be determined but our results imply that KZNFs have become widely integrated into neuronal gene regulatory networks. Our analysis predicts that gene expression networks have been repeatedly innovated throughout primate evolution, continuously gaining new layers of gene regulation mediated by both TEs and KZNFs in our genome.

Key Words: Krab Zinc Finger Proteins, Transposable elements, Evolutionary arms race, co-option
Chapter 2

Introduction

The family of KRAB zinc finger genes (KZNFs) is a large and rapidly evolving gene family recently shown to be involved in repression of transposable elements in mammalian genomes. The presence of a large number of young, species-specific KZNFs alongside ancient, highly conserved KZNFs suggests that this gene family could have been guarding our genomes for a long time. For some KZNFs, the structural adaptations they underwent throughout evolution suggests that their main role is repressing the retrotransposition activity of TEs in an attempt to stop them from spreading throughout the genome. These KZNFs and the TEs they evolved to repress, show clear signs of an evolutionary arms race that played out over the course of millions of years (Jacobs et al., 2014; Imbeault et al., 2017). However, because many KZNFs are still retained in our genome long after the invasion of the TE they evolved to repress has ceased, it is likely that KZNFs are co-opted for other regulatory roles. Indeed, for some KZNFs their acquired ability to recognize TEs seems to have been repurposed for tissue-specific regulation of TE-derived gene regulatory elements (Pontis et al., 2019). It was recently shown that KZNFs bind to TEs in early embryonic stages and use TEs to modulate nearby gene expression. Importantly, it was shown that the regulatory influence of KZNFs on TEs becomes evident again at later stages of human brain development (Pontis et al., 2019). Whether KZNFs also have other regulatory functions independent of TEs remains elusive. Here, we find that evolutionary changes in KZNFs provide them not only with the ability to recognize and repress TEs, but also enabled some KZNFs to bind to gene promoters. We reanalysed KZNF binding site data (Najafabadi et al., 2015; Imbeault et al., 2017; Schmitges et al., 2016) and found that a subset of KZNFs also bind to a large number of human gene promoters. The observation that in the human brain, the expression of these KZNF genes shows widespread correlations with the expression of their target genes suggests that alongside their involvement in repressing and modulating the activity of TEs, some KZNF genes integrate into existing gene expression networks and act as classical TFs to establish novel gene regulatory networks.
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns

Results

KRAB-ZNFs bind to a large number of gene promoters

Previous ChIP-seq and ChIP-exo analyses have revealed widespread binding of KZNFs to TEs. To a lesser extent, some KZNFs were reported to bind to non-TE sequences as well, but the scope and the implications of this phenomenon were not thoroughly investigated. To explore the non-TE binding sites of KZNFs, we reanalysed the MACS-peak data for 197 KZNFs and focused on the binding sites that overlapped with gene promoters. 51 of the KZNFs analysed were found to bind to at least 50 gene promoters (Figure 1A). Even though the majority (45/51) of the KZNFs bind to a few hundred or fewer gene promoters genome-wide, some KZNFs (ZNF202, ZNF534, ZNF519, ZNF263, ZNF282, ZNF257) showed binding to more than a thousand promoters. Notably, there was not a big overlap of the set of promoters bound by each of the KZNFs, in line with the highly divergent DNA binding domains of KZNFs: In total, 8888 genes showed binding of at least one KZNF in their promoter region, which correlates to ~45% of the genes in the human genome. The ability of the KZNFs to bind to gene promoters was not restricted to very old KZNF genes, as almost half of the promoter-binding KZNFs were found to be primate-specific (Thomas & Schneider 2011). Taking this into account, ~26% of human gene promoters have a binding site for at least one primate-specific KZNF.

The ChIP-seq data used in this analysis was generated using ectopic expression of KZNFs, raising the possibility that the ectopically expressed KZNF-GFP and KZNF-HA fusion proteins may bind to sites where the endogenous KZNF does not bind under physiological conditions. To assess if the observed promoter binding of KZNFs is a general artefact of the methodology used, we compared ChIP-seq signals of endogenous KZNFs with signals obtained from overexpression of these proteins (Figure 1B). ChIP data from overexpressed ZNF675 and ZNF141 was compared to ChIP-seq data generated with an antibody that recognizes endogenous ZNF93 (ab104878; Jacobs et al., 2014), but also targets a small number of other ZNFs including ZNF254, ZNF675 and ZNF141. In this comparison, we observe endogenous KZNF binding to the promoter regions of genes such as CALCRL and PLPP5 on the same location as where binding was observed for ectopically expressed ZNF141 and ZNF675 (Imbeault et al., 2017). This validates that the binding pattern of ectopically expressed KZNFs can be used as a representation of the potential binding pattern of the endogenous KZNFs. This is further supported by previous observations where binding of ectopically expressed KZNFs was compared to endogenous TRIM28/KAP1 (Imbeault et al., 2017). These data also confirmed that the binding pattern of ectopically
expressed KZNFs closely resembles the binding pattern of endogenous KAP1, which is recruited to KZNFs through their KRAB domains (Imbeault et al., 2017). Although this does not rule out that some of the KZNF binding can still be an artefact of ectopic expression, our analysis validates that promoter binding by endogenous ZNFs does occur under normal cellular conditions with endogenous levels of KZNF expression.

**KRAB-ZNFs show different expression patterns across time and space in the brain**

It was previously shown that KZNFs are highly expressed in the adult human brain (Imbeault et al., 2017). To determine the expression dynamics of KZNFs throughout pre- and postnatal human brain development, we plotted the BrainSpan RNA-seq expression levels (Miller et al., 2014) of each of the 51 KZNFs across various neurodevelopmental timepoints (Figure 2). Expression-based clustering shows some KZNFs have a uniform low or high expression across developmental time and some KZNFs are expressed at high levels in early development, followed by decreased expression in adult stages. Comparing the expression patterns between different brain regions revealed that the dynamics of KZNF expression is brain region specific. The developmental gene expression pattern of KZNFs in whole brain (Figure 2) was most similar to dorsolateral prefrontal cortex (DFC) but diverged significantly in other brain regions, such as the cerebellum (CBC) (Supplemental figure 1).
Figure 1. Analysis of KZNF ChIP-seq data at gene promoters. 
A) The number of gene promoters bound by KZNFs. Only KZNFs with >50 promoter targets are shown. Dark blue bars denote primate-specific KZNFs. Light blue bars indicate evolutionary older, mammalian KZNFs. B) ChIP-seq data for experiments using antibodies for KZNF-GFP fusion proteins and endogenous KZNFs showing peaks at gene promoters. Each window was set to the same scale per KZNF dataset.
Figure 2. Heatmap showing whole-brain expression over time of all KZFPs that bind >50 gene promoters.
Brainspan expression values in RKPM, low expression in green, high expression denoted by red. Ages: postconceptional weeks (pcw), months (mos), years (yrs)
Expression correlation of KRAB-ZNFs and target genes

We next addressed whether we could find correlations between the expression of the KZNFs and their target genes (Figure 3). For each of the selected 51 KZNFs, expression correlation values of all genes expressed in the brain was downloaded from the BrainSpan database (Miller et al., 2014). The distribution of these correlation values was plotted in a histogram and, as expected, a normal distribution of the expression correlation values, clustered around 0, was observed (Figure 3A; red histogram). Next, the correlation values for KZNF-target genes in relation to the expression of their respective KZNF were extracted and plotted in a histogram. If no regulatory effect is present of the KZNF on its target promoters a normal distribution of the expression correlation values would be expected, as previously observed for all genes. Remarkably, for many KZNFs we investigated, the expression correlation values of the KZNF target genes were not normally distributed, and the KZNF target genes often showed strong correlation or anti-correlation to the expression pattern of the KZNF gene (Figure 3A; blue histogram). As a control, we plotted an equal amount of randomly selected genes, and compared it to the distribution of all brain-expressed genes (Figure 3B). Randomly selected genes show a normal distribution that closely matches the distribution of correlation values of all genes when correlated to the expression of the KZNF. Based on the distribution patterns of the correlation values, the KZNFs could be grouped in four classes (Figure 3C): class I had distributions skewed towards positive correlation values, class II had uniform distributions, class III had distributions skewed towards negative correlation values, and class IV followed the ‘normal’ distribution of the values for all genes (Figure 3C; supplemental Figure 2). When compared using a Kruskal-Wallis test, measures of skew and kurtosis were significantly different between these groups (p-values < 0.001). These analyses support a regulatory relationship between 37/51 (72%) of the KZNFs and their target genes.
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns

Figure 3. Overview of KZNF - target gene correlation. (page 36)
A) Distribution of ZNF519 gene correlation values. Blue histogram indicates ZNF519 correlation with target genes. Red histogram indicates ZNF519 correlation with all genes. B) Distribution of ZNF519 gene correlation values using a random set of target genes (in grey) instead. C) Heatmap showing the distribution of expression correlation values of genes bound by the ZNFs. Each square represents the percentage of target genes that fall within the range of correlation values covered by each bin. KZNFs clustered using the heatmap.2 dendrogram function, classes were defined using this unbiased clustering. On the right, representative histograms for each cluster showing in blue the distribution of target gene expression correlation values and in red the values of all genes relative to the zinc finger. Black - red gradient indicates increasing percentage of genes per bin. Max value 20% (red), lowest value 0% (black).

Based on gene expression correlations, KZNFs act as regular Transcription factors

To compare the patterns of correlations between KZNFs and their target genes to other TFs and their target genes, the expression correlation analysis was repeated on a set of 12 well documented TFs, part of the ENCODE consortium dataset (ENCODE Project Consortium, 2012) (Figure 4). The distribution patterns of the correlation values for these TFs and their target genes showed a remarkably similar pattern to those observed for KZNFs. The described role of the TEs as activators or repressors does not seem to affect the distribution patterns. The known TFs also form similar clusters based on the distribution of the expression correlation values of the genes to which they bind. When choosing a random set of gene-targets, we again observed a distribution mirroring the distribution of all genes instead. The similar findings for the KZNF and TF datasets indicates that certain KZNFs may have been repurposed to act as traditional transcription factors.
Figure 4. Heatmap showing the distribution of expression correlation values of genes bound by 12 documented transcription factors. Generated with the same method as figure 3. To the right, representative histograms for each of the clusters. Black – red gradient indicates increasing percentage of genes per bin. Max value 20% (red), lowest value 0% (black)

The correlation between KZNF and target genes is tissue dependent

Our analysis revealed brain-region specific expression of KZNFs during development. We therefore investigated whether there is region-specific regulatory influence of KZNFs on their target gene promoters (Figure 5). For this analysis we selected 10 KZNFs that bind to the highest number of gene promoters. Similar to the differential gene expression dynamics of KZNFs in different brain regions, the expression correlation distributions are not consistent across the different brain regions. In the cerebellum values seem to
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns

Figure 5. Heatmaps showing the distribution of expression correlation values for each of the top 10 promoter-binding KZNFs across different brain regions. (cont. p38)
Data from the Dorsolateral Frontal Cortex (DFC), Cerebellum (CBC), Thalamus (THM) and Striatum (STR). ZNF479 showed no expression or correlation values in the striatum and is depicted as a black line here. Order of ZNFs denoted by the clustering in the DFC. Black - red gradient indicates increasing percentage of genes per bin. Max value 20% (red), lowest value 0% (black)

Cluster around 0, suggesting that KZNFs do not have a major influence on gene expression in this region. The highest divergence from the normal distribution of all genes was observed for DFC, which indicates that in this tissue KZNFs have the highest influence on the expression of their target genes (Figure 6). Overall, our data suggests that the binding of KZNFs to gene promoters is associated with regulatory effects, and shows a brain-region dependent influence on the expression of target genes.
Figure 6. Gene correlation histograms for ZNF257. 
A), B) showing distributions of expression correlation values of target genes (A) and a set of random genes (B) in the whole brain compared to all genes. All genes are shown in red, target genes in blue, and random genes in grey. (C-F) showing the distribution of expression correlation values of target genes vs all genes in the different brain areas (C) dorsolateral prefrontal cortex, (D) cerebellar cortex, (E) thalamus, (F) striatum. For these regions, correlation values of 0.000 were removed due to the large number of unexpressed genes in specific brain regions.
Discussion

Our analysis of KZNF binding sites suggest that gene promoters have become regulatory targets of KZNFs thereby forming a new layer of gene regulation as a consequence of the evolutionary arms race between KZNFs and TEs. This is supported by previous studies showing links between KZNFs and gene expression (Chen et al., 2019, Ecco et al., 2016, Yang et al., 2017). The question arises why and how these gene promoters became bound by the KZNFs. It was previously shown that some KZNFs and the transposable elements they repress are locked in an evolutionary arms race, each placing selective pressure on the other to evolve and fixate mutations in their DNA sequences. It is not unlikely that during the phase of structural evolution of the KZNF, when optimizations occur in the DNA binding domains, the KZNF evolves to recognize and bind other non-TE sequences by chance (Figure 7). Another scenario is that some KZNFs by default recognize GC-rich sequences, a feature often observed in both TEs and gene promoters (Lander et al., 2001). There may be no possibility for a KZNF to bind and repress a GC-rich region in a TE without also binding and repressing GC-rich regions in gene promoters. In a third scenario, the binding of KZNFs to gene promoters may be explained by the strategy viruses use to maximize the production of viral components by the host cells. One of the most efficient ways of establishing this is having an LTR region which contains sequences resembling the promoters of highly expressed genes. For the host, in its battle to control the retroviral-activity and/or retrotransposon invasion, the most efficient strategy would be to repress the LTR region responsible for expression of these retroviral components. The consequence is that structural evolution of KZNFs to repress TEs results in the KZNF repressing the gene promoters that the original LTR was mimicking (Figure 7). What is relevant to note is that we observed many promoters bound by more than one KZNF. However, the binding sites of multiple KZNFs within the same promoter are usually not overlapping, showing a clear sequence specificity of KZNFs even in the repetitive parts of TEs and promoters. All of the scenarios above provide a good potential explanation for the fixation of KZNF genes after the TEs they evolved to repress have become inactive and neutralized. As such, each KZNF gene may have evolved via one of these scenarios, which may play on KZNF structural evolution as complementing evolutionary forces. Importantly, the non-TE mediated binding of gene promoters by some KZNFs is complemented by other KZNFs that predominantly regulate gene expression by binding to TE-derived regulatory elements (Pontis et al., 2019). Together they provide a picture of massive co-option of TEs, and also KZNFs, as two heavily interdependent layers of gene regulation.
The role of KZNFs as gene regulators is supported by our analysis of expression correlation data. The expression of both KZNFs and TFs is correlated with the expression of the genes to which they bind. Out of the 10 KZNFs that bind the highest number of gene promoter regions, 7 have a correlation distribution skewed to the positive side, where the expression of bound genes increases with the expression of the KZNF. This is at odds with the well documented, repressive role of KZNFs but could be explained by the phenomena where KZNFs, especially ancient ones, lost their ability to recruit TRIM28 (Imbeault et al., 2017). Indeed, a comparison of the promoter binding sites of the top 10 KZNFs with the binding sites of KAP1/TRIM28 show very little overlap (Supplemental table 1). It remains unknown why KZNFs lose the ability to recruit TRIM28 and what happens after they lose this ability, but one possibility is that loss of TRIM28 recruiting capacity may provide a mechanism in which KZNFs can be transformed into positive regulators of gene expression. This yields additional flexibility for KZNFs to integrate into gene-regulatory networks.

For the gene promoters bound by KZNFs that retain the ability to recruit TRIM28 another question arises: What is the impact of this repressive machinery on gene expression? If important neurodevelopmental genes are bound, this could be negative for the organism. This, in turn could lead to a selective pressure against binding of gene promoters by KZNFs and could explain why the majority of KZNFs bind <200 gene promoters. In combination with the KZNFs that lost their KRAB domain and repressive capabilities, this could exemplify how KZNFs form a complex gene regulatory network in the brain, consisting of a combination of repressive and promoting effects. The combined binding of multiple brain-expressed KZNFs to gene promoters may have subtle but significant effects on transcript levels of genes crucial to brain development. Taken together, our analysis suggests that both TEs and KZNFs could have contributed new primate-specific layers of gene regulation to our genome. This underlines how the intimate relation between TEs and KZNFs has the potential to repeatedly cause innovations in gene expression networks throughout primate evolution.

Figure 7. Model showing a mechanism by which gene promoters may have become bound by KZFPs
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns

Materials & Methods

Analysis of MACS-peak data and KZNF binding sites

The ChIP-seq and ChIP-exo data used for this experiment was generated by Schmitges et al. (2016) (NCBI GEO database accession number GSE76496), Imbeault et al. (2017) (NCBI GEO database accession number GSE78099), Najafabadi et al. (2015) (NCBI GEO database accession number GSE52523) and Jacobs et al. (2014) (NCBI GEO database accession number GSE60210). To visualize the data, raw fastq files were imported to the Galaxy US or EU servers (Afgan et al., 2018). Reads were processed, adaptor and illumina-specific sequences were removed using Trimmomatic (Galaxy v0.36.5), and mapped to the human genome (assembly GRCh37/hg19; Lander et al., 2001) using Bowtie2 (Galaxy v2.3.4.2) (Langmead & Salzberg, 2012) with single-end, very sensitive end-to-end settings. BigWig files were generated using the bamCoverage tool (Galaxy v3.0.1.0). These were uploaded to the UCSC Genome Browser (Kent et al., 2002) to manually confirm the reads were in concordance with the peak calling data.

A cut-off of 500 for the p-value score (10xlog10(pvalue)) was decided after visual observation of the MACS-peaks using the UCSC genome browser. The script Match_final.py was used to compare remaining peaks with the location of gene promoter regions acquired from UCSC data hubs (Raney et al., 2014). The region 1000bp downstream and 5000bp upstream of a gene transcription start site was taken and any gene promoter that overlapped with at least 50% of a MACS-peak was recorded using the match_prom() function. Further details on the function of all scripts can be found in the supplemental python files.

Analysis of KZNF expression data

KZNF expression data was obtained from the BrainSpan Gene Expression tool (Miller et al., 2014). The function data_maker(brainarea) was used to extract the expression data for all the KZNFs with >50 binding sites on gene promoter regions. Heatmaps for whole brain expression as well as expression in specific brain areas were generated with R (R Development Core Team, 2008) using the heatmap.2 function from the library gplots. For the specific brain regions, row order was determined by the clustering of the whole brain heatmap.
Chapter 2

Analysis of KZNF expression correlation data

The api.py script contains the functions used in the following steps. Expression correlation values for all genes in the whole brain in relation to each zinc finger were downloaded from www.brainspan.org (2015 Allen Institute for Brain Science. Allen Brain Atlas API. Available from: brain-map.org/api/index.html) using their API. The correlation values for the genes directly bound by the KZNFs were extracted for each KZNF and the distribution of these correlation values was visualised in a histogram alongside the distribution of the expression correlation values of all genes relative to that KZNF. Histograms were made using Matplotlib (Hunter, 2007). For the histograms for specific brain regions correlation values of 0.000 were removed.

Histograms were also generated showing the distribution of a random set of gene expression correlation values next to the distribution of expression values of all genes. The function rand_check_hist() selects at random an equal number of genes to the number of genes bound by the KZNF in question.

Heatmaps showing the distribution of the expression correlation values of the target genes of every KZNF in the whole brain and specific regions were generated using the function heatmap_exp_cor_dis() and the heatmap.2 function in R. The function assigns the segments of the heatmap based on the percentage of target genes that have a correlation value that falls within the range determined on the x-axis. If the bins 0.8 to 1.0 contained no values, they were removed. Row order for the specific tissues was based on the clustering for the dorsolateral prefrontal cortex (DFC). The heatmap.2 clustering algorithm was used to group the KZNFs into 4 categories. The skew and kurtosis of each distribution was calculated using the R package e1071. The skew and kurtosis of the groups were then tested for statistical difference using a Kruskal-Wallis test in R.

Analysis of transcription factor data

Binding site data for TFs was obtained from the ENCODE database (ENCODE Project Consortium, 2012). The functions match_prom_trans(tscf) and match_symbols_trans(tscf) in the Match_final.py script were used to generate lists of target gene promoters. This was then subjected to the same workflow as the ZNF gene expression correlations data to generate correlation histograms and heatmaps using the slightly altered heatmap_tscf_correlation() function.
Author Contributions

Acknowledgements
We thank Michael Imbeault and members of the Jacobs lab for helpful discussions. This project was supported by ERC-2016-StG-716035 (F.M.J.J.) and HFSP CDA00030/2016C (F.M.J.J.)
Chapter 2

References


Kent, W. J. et al. The human genome browser at UCSC. Genome Res. 12, 996–1006 (2002).


Supplemental figure 1. KZNF expression during development in different brain regions. Showing data from the dorsolateral prefrontal cortex (DFC), cerebellar cortex (CBC), Thalamus (THM), striatum (STR). Heatmaps were generated using all data available. Note that not all time points are equally represented for each brain region.
Supplemental Figure 2. Histograms showing the distribution of expression correlation values between the top 10 promoter-binding KZNFs. For each KZNF, correlation distributions of all genes (red) and target genes are shown (blue). Histograms grouped by clusters based on distribution patterns.
Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns

Supplemental Figure 3. Histograms showing the distribution of expression correlation values for multiple transcription factors. For each transcription factor, correlation distributions with all genes (red) and target genes are shown (blue).
## Chapter 2

<table>
<thead>
<tr>
<th>ZNF</th>
<th>Primate specific</th>
<th>Traced back to: (Thomas &amp; Schneider 2011)</th>
<th># Promoter binding sites that overlap with KAP1 binding sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZNF202</td>
<td>N</td>
<td>Marsupial</td>
<td>35</td>
</tr>
<tr>
<td>ZNF257</td>
<td>Y</td>
<td>Simian</td>
<td>6</td>
</tr>
<tr>
<td>ZNF263</td>
<td>N</td>
<td>Marsupial</td>
<td>18</td>
</tr>
<tr>
<td>ZNF282</td>
<td>N</td>
<td>Marsupial</td>
<td>18</td>
</tr>
<tr>
<td>ZNF441</td>
<td>Y</td>
<td>Simian</td>
<td>4</td>
</tr>
<tr>
<td>ZNF468</td>
<td>Y</td>
<td>Catarrhine</td>
<td>12</td>
</tr>
<tr>
<td>ZNF479</td>
<td>N</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>ZNF519</td>
<td>Y</td>
<td>Hominoid</td>
<td>30</td>
</tr>
<tr>
<td>ZNF534</td>
<td>Y</td>
<td>Catarrhine</td>
<td>32</td>
</tr>
<tr>
<td>ZNF783</td>
<td>N</td>
<td>Marsupial</td>
<td>5</td>
</tr>
</tbody>
</table>

**Supplemental Table 1.** The top 10 binding KZNFs, their age as defined by Thomas & Schneider (2011) and the number of binding sites that overlap with both gene promoters and KAP1 binding sites (Imbeault et al., 2017)