

This is a repository copy of *Transcription factor clusters regulate genes in eukaryotic cells*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/121854/>

Version: Accepted Version

Article:

Wollman, Adam orcid.org/0000-0002-5501-8131, Shashkova, Sviatlana, Hedlund, Erik Goran et al. (3 more authors) (2017) Transcription factor clusters regulate genes in eukaryotic cells. *eLife*. e27451. ISSN 2050-084X

<https://doi.org/10.7554/eLife.27451>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

25 **Abstract**

26 Transcription is regulated through binding factors to gene promoters to activate or repress
27 expression, however, the mechanisms by which factors find targets remain unclear. Using
28 single-molecule fluorescence microscopy, we determined *in vivo* stoichiometry and
29 spatiotemporal dynamics of a GFP tagged repressor, Mig1, from a paradigm signaling
30 pathway of *Saccharomyces cerevisiae*. We find the repressor operates in clusters, which upon
31 extracellular signal detection, translocate from the cytoplasm, bind to nuclear targets and
32 turnover. Simulations of Mig1 configuration within a 3D yeast genome model combined with
33 a promoter-specific, fluorescent translation reporter confirmed clusters are the functional unit
34 of gene regulation. *In vitro* and structural analysis on reconstituted Mig1 suggests that
35 clusters are stabilized by depletion forces between intrinsically disordered sequences. We
36 observed similar clusters of a co-regulatory activator from a different pathway, supporting a
37 generalized cluster model for transcription factors that reduces promoter search times through
38 intersegment transfer while stabilizing gene expression.

40 **Introduction**

41 Cells respond to their environment through gene regulation involving protein transcription
42 factors. These proteins bind to DNA targets of a few tens of base pairs (bp) length inside
43 ~500-1,000bp promoter sequences to repress/activate expression, involving single (1) and
44 multiple (2) factors, resulting in the regulation of target genes. The mechanism for finding
45 targets in a genome ~six orders of magnitude larger is unclear since free diffusion followed
46 by capture is too slow to account for observed search times (3). Target finding may involve
47 heterogeneous mobility including nucleoplasmic diffusion, sliding and hops along DNA up to
48 ~150bp, and even longer jumps separated by hundreds of bp called intersegment transfer (4–
49 6).

50 In eukaryotes, factor localization is dynamic between nucleus and cytoplasm (7).
51 Although target binding sites in some cases are known to cluster in hotspots (8) the
52 assumption has been that factors themselves do not function in clusters but as single
53 molecules. Realistic simulations of diffusion and binding in the complex milieu of nuclei
54 suggest a role for multivalent factors to facilitate intersegment transfer by enabling DNA
55 segments to be connected by a single factor (9).

56 The use of single-molecule fluorescence microscopy to monitor factor localization in
57 live cells has resulted in functional insight into gene regulation (10). Fluorescent protein
58 reporters, in particular, have revealed complexities in mobility and kinetics in bacterial (11)
59 and mammalian cells (12) suggesting a revised view of target finding (4).

60 Key features of gene regulation in eukaryotes are exemplified by glucose sensing in
61 budding yeast, *Saccharomyces cerevisiae*. Here, regulation is achieved by factors which
62 include the Mig1 repressor, a Zn finger DNA binding protein (13) that acts on targets
63 including *GAL* genes (14). Mig1 is known to localize to the nucleus in response to increasing
64 extracellular glucose (15), correlated to its dephosphorylation (16, 17). Glucose sensing is
65 particularly valuable for probing gene regulation since the activation status of factors such as
66 Mig1 can be controlled reproducibly by varying extracellular glucose. Genetic manipulation
67 of the regulatory machinery is also tractable, enabling native gene labeling with fluorescent
68 reporters for functioning imaging studies.

69 We sought to explore functional spatiotemporal dynamics and kinetics of gene
70 regulation in live *S. cerevisiae* cells using its glucose sensing pathway as a model for signal
71 transduction. We used single-molecule fluorescence microscopy to track functional
72 transcription factors with millisecond sampling to match the mobility of individual

73 molecules. We were able to quantify composition and dynamics of Mig1 under physiological
74 and perturbed conditions which affected its possible phosphorylation state. Similarly, we
75 performed experiments on a protein called Msn2, which functions as an activator for some of
76 Mig1 target genes (18) but controlled by a different pathway. By modifying the microscope
77 we were also able to determine turnover kinetics of transcription factors at their nuclear
78 targets.

79 The results, coupled to models we developed using chromosome structure analysis,
80 indicated unexpectedly that the functional component which binds to promoter targets
81 operates as a cluster of transcription factor molecules with stoichiometries of ~6-9 molecules.
82 We speculated that these functional clusters in live cells were stabilized through interactions
83 of intrinsically disordered sequences facilitated through cellular depletion forces. We were
84 able to mimic those depletion forces in *in vitro* single-molecule and circular dichroism
85 experiments using a molecular crowding agent. Our novel discovery of factor clustering has a
86 clear functional role in facilitating factors finding their binding sites through intersegment
87 transfer, as borne out by simulations of multivalent factors (9); this addresses a long-standing
88 question of how transcription factors efficiently find their targets. This clustering also
89 functions to reduce off rates from targets compared to simpler monomer binding. This effect
90 improves robustness against false positive detection of extracellular chemical signals, similar
91 to observations for the monomeric but multivalent bacterial LacI repressor (4). Our findings
92 potentially reveal an alternative eukaryotic cell strategy for gene regulation but using an
93 entirely different structural mechanism.

94

95 **Results**

96 **Single-molecule imaging reveals *in vivo* clusters of functional Mig1**

97 To explore the mechanisms of transcription factor targeting we used millisecond Slimfield
98 single-molecule fluorescence imaging (19–21) on live *S. cerevisiae* cells (Figure. 1A and
99 Figure. 1 – Figure Supplement 1). We prepared a genomically encoded green fluorescent
100 protein (GFP) reporter for Mig1 (Table 1). To enable nucleus and cell body identification we
101 employed mCherry on the RNA binding nuclear protein Nrd1. We measured cell doubling
102 times and expression to be the same within experimental error as the parental strain
103 containing no fluorescent protein (Figure. 1 – Figure Supplement 2A). We optimized
104 Slimfield for single-molecule detection sensitivity with an *in vitro* imaging assay of surface-
105 immobilized purified GFP (22) indicating a brightness for single GFP molecules of
106 ~5,000 counts on our camera detector (Figure. 1 – Figure Supplement 2B). To determine any
107 fluorescent protein maturation effects we performed cell photobleaching while expression of
108 any additional fluorescent protein was suppressed by antibiotics, and measured subsequent
109 recovery of cellular fluorescence <15% for fluorescent protein components, corrected for any
110 native autofluorescence, over the timescale of imaging experiments (Figure. 1 – Figure
111 Supplement 2C and D).

112 Under depleted (0%)/elevated (4%) extracellular glucose (-/+), we measured
113 cytoplasmic and nuclear Mig1 localization bias respectively, as reported previously (15),
114 visible in individual cells by rapid microfluidic exchange of extracellular fluid (Figure. 1B),
115 with high cell-cell variability (Figure. 1B middle panel). However, our ultrasensitive imaging
116 resolved two novel components under both conditions consistent with a diffuse monomer
117 pool and distinct multimeric foci which could be tracked up to several hundred milliseconds
118 (Figure. 1C and Figure Supplement 3; Video 1 and 2). We wondered if the presence of foci
119 was an artifact due to GFP oligomerization. To discourage artifactual aggregation we

120 performed a control using another type of GFP containing an A206K mutation (denoted
121 GFPmut3 or mGFP) known to inhibit oligomerization (23). However, both *in vitro*
122 experiments using purified GFP and mGFP (Figure. 1 – Figure Supplement 2B) and live cell
123 experiments at *glucose* (-/+) (Figure. 1 – Figure Supplement 2E and F) indicated no
124 significant difference to foci brightness values (Student's t-test, $p=0.67$). We also developed a
125 genomically encoded Mig1 reporter using green-red photoswitchable fluorescent protein
126 mEos2 (24). Super-resolution stochastic optical reconstruction microscopy (STORM) from
127 hundreds of individual photoactivated tracks indicated the presence of foci, clearly present in
128 nuclei hotspots in live cells at *glucose* (+) (Figure. 1 – Figure Supplement 1). These results
129 strongly argue that foci formation is not dependent on hypothetical fluorescent protein
130 oligomerization.

131 We implemented nanoscale tracking based on automated foci detection which
132 combined iterative Gaussian masking and fitting to foci pixel intensity distributions to
133 determine the spatial localization to a lateral precision of 40nm (25, 26). Tracking was
134 coupled to stoichiometry analysis using single GFP photobleaching of foci tracks (22) and
135 single cell copy number quantification (27). These methods enabled us to objectively
136 quantify the number of Mig1 molecules associated with each foci, its effective microscopic
137 diffusion coefficient D and spatiotemporal dynamics in regards to its location in the
138 cytoplasm, nucleus or translocating across the nuclear envelope, as well as the copy number
139 of Mig1 molecules associated with each subcellular region and in each cell as a whole. These
140 analyses indicated ~850-1,300 Mig1 total molecules per cell, dependent on extracellular
141 glucose. Quantitative PCR and previous work suggest a higher Mig1 copy number at *glucose*
142 (-) (27) (Figure. 1D; Table 2 and 3).

143 At *glucose* (-) we measured a mean ~950 Mig1 molecules per cell in the cytoplasmic
144 pool (Figure. 1D) and 30-50 multimeric foci in total per cell, based on interpolating the
145 observed number of foci in the microscope's known depth of field over the entirety of the cell
146 volume. These foci had a mean stoichiometry of 6-9 molecules and mean D of 1-2 $\mu\text{m}^2/\text{s}$,
147 extending as high as 6 $\mu\text{m}^2/\text{s}$. In nuclei, the mean foci stoichiometry and D was the same as
148 the cytoplasm to within experimental error (Student's t-test, $p>0.05$, $p=0.99$ and $p=0.83$),
149 with a similar concentration. Trans-nuclear foci, those entering /leaving the nucleus during
150 observed tracking, also had the same mean stoichiometry and D to cytoplasmic values to
151 within experimental error ($p>0.05$, $p=0.60$ and $p=0.79$). However, at *glucose* (+) we
152 measured a considerable increase in the proportion of nuclear foci compared to *glucose* (-),
153 with up to 8 foci per nucleus of mean apparent stoichiometry 24-28 molecules, but D lower
154 by a factor of 2, and 0-3 cytoplasmic/trans-nuclear foci per cell (Figure. 2A and 2B and
155 Figure Supplement 3).

156

157 **Mig1 cluster localization is dependent on phosphorylation status**

158 To understand how Mig1 clustering was affected by its phosphorylation we deleted the *SNF1*
159 gene which encodes the Mig1-upstream kinase, Snf1, a key regulator of Mig1
160 phosphorylation. Under Slimfield imaging this strain indicated Mig1 clusters with similar
161 stoichiometry and D as for the wild type strain at *glucose* (+), but with a significant
162 insensitivity to depleting extracellular glucose (Figure. 1 – Figure Supplement 1, Figure. 2 –
163 Figure Supplement 1A and B). We also used a yeast strain in which the kinase activity of
164 Snf1 could be controllably inhibited biochemically by addition of cell permeable PP1 analog
165 1NM-PP1. Slimfield imaging indicated similar results in terms of the presence of Mig1
166 clusters, their stoichiometry and D , but again showing a marked insensitivity towards
167 depleted extracellular glucose indistinguishable from the wild type *glucose* (+) phenotype

168 (Figure. 1 - Figure Supplement 1, Figure. 2 – Figure Supplement 1C, Figure. 2 – Figure
169 Supplement 2,3 and Table 4). We also tested a strain containing Mig1 with four serine
170 phosphorylation sites (Ser222, 278, 311 and 381) mutated to alanine, which were shown to
171 affect Mig1 localization and phosphorylation dependence on extracellular glucose (28).
172 Slimfield showed the same pattern of localization as the *SNF1* deletion while retaining the
173 presence of Mig1 clusters (Figure. 2 – Figure Supplement 1D and E). These results suggest
174 that Mig1 phosphorylation does not affect its ability to form clusters, but does alter their
175 localization bias between nucleus and cytoplasm.

176

177 **Cytoplasmic Mig1 is mobile but nuclear Mig1 has mobile and immobile states**

178 The dynamics of Mig1 between cytoplasm and nucleus is critically important to its role in
179 gene regulation. We therefore interrogated tracked foci mobility. We quantified cumulative
180 distribution functions (CDFs) for all nuclear and cytoplasmic tracks (12). A CDF signifies the
181 probability that foci will move a certain distance from their starting point as a function of
182 time while tracked. Here, we analyzed only the first displacement of each track to avoid bias
183 toward slowly moving foci (12). A mixed mobility population can be modeled as the
184 weighted sum of multiple CDFs characterized by different *D*. Cytoplasmic foci at *glucose*
185 (+/-), and nuclear foci at *glucose* (-), were consistent with just a single mobile population
186 (Figure. 3 – Figure Supplement 1) whose *D* of 1-2 $\mu\text{m}^2/\text{s}$ was consistent with earlier
187 observations. However, nuclear foci at *glucose* (+) indicated a mixture of mobile and
188 immobile components (Figure. 3A). These results, substantiated by fitting two Gamma
189 functions to the distribution of estimated *D* (29) for *glucose* (+) nuclear foci (Figure. 3A,
190 inset), indicate 20-30% of nuclear foci are immobile, consistent with a DNA-bound state.
191 Mean square displacement analysis of foci tracks sorted by stoichiometry indicated Brownian
192 diffusion over short timescales of a few tens of ms but increasingly anomalous diffusion over
193 longer timescales >30ms (Figure. 3B). These results are consistent with *glucose* (+) Mig1
194 diffusion being impacted by interactions with nuclear structures, similar to that reported for
195 other transcription factors (30). Here however this interaction is dependent on extracellular
196 glucose despite Mig1 requiring a pathway of proteins to detect it, unlike the more direct
197 detection mechanism of the prokaryotic *lac* repressor. A strain in which mCherry labeled
198 Mig1 had its Zn finger deleted ($\Delta\text{aa}36-91$) (18) indicated no significant immobile cluster
199 population at *glucose* (+/-) with CDF analysis (Figure. 3 – Figure Supplement 1). We
200 conclude that Mig1 clusters bind with a relatively high association constant to the DNA via
201 their Zn finger motif with direct glucose dependence.

202

203 **Mig1 nuclear translocation selectivity does not depend on glucose but is mediated by** 204 **interactions away from the nuclear envelope**

205 Due to the marked localization of Mig1 towards nucleus/cytoplasm at *glucose* (+/-)
206 respectively, we asked whether this spatial bias was due to selectivity initiated during
207 translocation at the nuclear envelope. By converting trans-nuclear tracks into coordinates
208 parallel and perpendicular to the measured nuclear envelope position, and synchronizing
209 origins to be the nuclear envelope crossing point, we could compare spatiotemporal dynamics
210 of different Mig1 clusters during translocation. A heat map of spatial distributions of
211 translocating clusters indicated a hotspot of comparable volume to that of structures of
212 budding yeast nuclear pore complexes (31) and accessory nuclear structures of cytoplasmic
213 nucleoporin filaments and nuclear basket (32), with some nuclear impairment to mobility
214 consistent with restrained mobility (Figure. 3C). We observed a dwell in cluster translocation

215 across the 30-40nm width of the nuclear envelope (Figure. 3D). At *glucose* (+) the proportion
216 of detected trans-nuclear foci was significantly higher compared to *glucose* (-), consistent
217 with Mig1's role to repress genes. The distribution of dwell times could be fitted using a
218 single exponential function with ~10ms time constant similar to previous estimates for
219 transport factors (33). However, although the relative proportion of trans-nuclear foci was
220 much lower at *glucose* (-) compared to *glucose* (+), the dwell time constant was found to be
221 insensitive to glucose (Figure. 3E). This insensitivity to extracellular chemical signal
222 demonstrates, surprisingly, that there is no direct selectivity on the basis of transcription
223 factor phosphorylation state by nuclear pore complexes themselves, suggesting that cargo
224 selectivity mechanisms of nuclear transport (34), as reported for a range of substrates, is blind
225 to the phosphorylation state. Coupled with our observation that Mig1 at *glucose* (-) does not
226 exhibit significant immobility in the nucleus and that Mig1 lacking the Zn finger still
227 accumulates in the nucleus at *glucose* (+) (Figure. 1 – Figure Supplement 1), this suggests
228 that Mig1 localization is driven by changes in Mig1 binding affinity to other proteins, within
229 e.g. the general corepressor complex (35), or outside the nucleus not involving the nuclear
230 pore complex.

231

232 **Mig1 nuclear foci bound to targets turn over slowly as whole clusters of ~7-9 molecules** 233 **in >100s**

234 To further understand the mechanisms of Mig1 binding/release during gene regulation we
235 sought to quantify kinetics of these events at Mig1 targets. By modifying our microscope we
236 could implement an independent focused laser path using the same laser source, enabling us
237 to use fluorescence recovery after photobleaching (FRAP) to probe nuclear Mig1 turnover.
238 The focused laser rapidly photobleached GFP content in cell nuclei in <200ms (Figure. 3F).
239 We could then monitor recovery of any fluorescence intensity by illuminating with
240 millisecond Slimfield stroboscopically as opposed to continuously to extend the observation
241 timescale to >1,000s. Using automated foci detection we could separate nuclear pool and foci
242 content at each time point for each cell. These analyses demonstrated measurable
243 fluorescence recovery for both components, which could be fitted by single exponentials
244 indicating fast recovery of pool at both *glucose* (-) and (+) with a time constant <5s but a
245 larger time constant at *glucose* (+) for nuclear foci >100s (Figure. 3G). Further analysis of
246 intensity levels at each time point revealed a stoichiometry periodicity in nuclear foci
247 recovery equivalent to 7-9 GFP molecules (Figure. 4 – Figure Supplement 1A), but no
248 obvious periodicity in stoichiometry measurable from pool recovery. An identical periodicity
249 within experimental error was measured from nuclear foci at *glucose* (+) in steady-state
250 (Figure. 4A). These periodicity values in Mig1 stoichiometry were consistent with earlier
251 observations for cytoplasmic and trans-nuclear clusters at *glucose* (+/-), and in the nucleus at
252 *glucose* (-), with mean stoichiometry ~7 molecules. These data taken as a whole clearly
253 suggest that molecular turnover at nuclear foci of Mig1 bound to its target genes occurs in
254 units of single clusters, as opposed to single Mig1 monomers.

255

256 **Mig1 clusters are spherical, a few tens of nm wide**

257 Our observations from stoichiometry, dynamics and kinetics, which supported the hypothesis
258 that functional clusters of Mig1 perform the role of gene regulation, also suggested an
259 obvious prediction in terms of the size of observed foci: the physical diameter of a multimeric
260 cluster should be larger than that of a single Mig1 monomer. We therefore sought to quantify
261 foci widths from Slimfield data by performing intensity profile analysis on background-

262 corrected pixel values over each foci image. The diameter was estimated from the measured
263 width corrected for motion blur due to particle diffusion in the sampling time of a single
264 image frame, minus that measured from single purified GFP molecules immobilized to the
265 coverslip surface in separate *in vitro* experiments. This analysis revealed diameters of
266 15-50nm at *glucose* (-), which showed an increase with foci stoichiometry S that could be
267 fitted with a power law dependence S^a (Figure. 4 – Figure Supplement 1B) with optimized
268 exponent a of 0.32 ± 0.06 (\pm SEM). Immuno-gold electron microscopy of fixed cells probed
269 with anti-GFP antibody confirmed the presence of GFP in 90nm cryosections with some
270 evidence of clusters containing up to 7 Mig1 molecules (Figure. 4 – Figure Supplement 1C),
271 however, the overall labeling efficiency was relatively low with sparse labelling in the
272 nucleus in particular, possibly as a consequence of probe inaccessibility, resulting in
273 relatively poor statistics. A heuristic tight packing model for GFP labeled Mig1 monomers in
274 each cluster predicts that, in the instance of an idealized spherical cluster, $a = 1/3$. Our data at
275 *glucose* (-) thus supports the hypothesis that Mig1 clusters have a spherical shape. For
276 nuclear foci at *glucose* (+) we measured larger apparent diameters and stoichiometries,
277 consistent with >1 individual Mig1 cluster being separated by less than our measured
278 ~200nm optical resolution limit. This observation agrees with earlier measurements of
279 stoichiometry periodicity for nuclear foci at *glucose* (+). In other words, that higher apparent
280 stoichiometry nuclear foci are consistent with multiple individual Mig1 clusters each
281 containing ~7 molecules separated by a nearest neighbor distance <200nm and so detected as
282 a single fluorescent foci.

283

284 **Clusters are stabilized by depletion forces**

285 Since we observed Mig1 clusters in live cells using Slimfield imaging we wondered if these
286 could be detected and further quantified using other methods. However, native gel
287 electrophoresis on extracts from Mig1-GFP cells (Figure. 4 – Figure Supplement 2A)
288 indicated a single stained band for Mig1, which was consistent with denaturing SDS-PAGE
289 combined with western blotting using recombinant Mig1-GFP, and protein extracts from the
290 parental cells which included no fluorescent reporter (Figure. 4 – Figure Supplement 2B and
291 C). Slimfield imaging on purified Mig1-GFP *in vitro* under identical imaging conditions for
292 live cells similarly indicated monomeric Mig1-GFP foci in addition to a small fraction of
293 brighter foci which were consistent with predicted random overlap of monomer images.
294 However, on addition of low molecular weight polyethylene glycol (PEG) at a concentration
295 known to mimic small molecule ‘depletion’ forces in live cells (36) we detected significant
296 numbers of multimeric foci (Figure. 4B and Figure Supplement 2D). Depletion is an entropic
297 derived attractive force which results from osmotic pressure between particles suspended in
298 solution that are separated by distances short enough to exclude other surrounding smaller
299 particles. Purified GFP alone under identical conditions showed no such effect (Figure. 4 –
300 Figure Supplement 2E). These results support a hypothesis that clusters are present in live
301 cells regardless of the concentration of extracellular glucose, which are stabilized by
302 depletion components that are lost during biochemical purification.

303

304 **Chromosome structure modeling supports a cluster binding hypothesis**

305 We speculated that Mig1 cluster-mediated gene regulation had testable predictions in regards
306 to the nuclear location of Mig1 at elevated extracellular glucose. We therefore developed
307 quantitative models to simulate the appearance of realistic images of genome-bound Mig1-
308 GFP at *glucose* (+). We used sequence analysis to infer locations of Mig1 binding sites in the

309 yeast genome, based on alignment matches to previously identified 17bp Mig1 target patterns
310 (37) which comprised conserved AT-rich 5bp and GC-rich 6bp sequences. In scanning the
311 entire *S. cerevisiae* genome we found >3,000 hits though only 112 matches for likely gene
312 regulatory sites located in promoter regions (Table 5). We mapped these candidate binding
313 sites onto specific 3D locations (Figure. 4C) obtained from a consensus structure for budding
314 yeast chromosomes based on 3C data (38). We generated simulated images, adding
315 experimentally realistic levels of signal and noise, and ran these synthetic data through the
316 same tracking software as for experimental data. We used identical algorithm parameters
317 throughout and compared these predictions to the measured experimental stoichiometry
318 distributions.

319 In the first instance we used these locations as coordinates for Mig1 monomer
320 binding, assuming that just a single Mig1 molecule binds to a target. Copy number analysis
321 of Slimfield data (Table 2) indicated a mean ~190 Mig1 molecules per cell associated with
322 nuclear foci, greater than the number of Mig1 binding sites in promoter regions. We assigned
323 112 molecules to target promoter binding sites, then assigned the remaining 78 molecules
324 randomly to non-specific DNA coordinates of the chromosomal structure. We included the
325 effects of different orientations of the chromosomal structure relative to the camera by
326 generating simulations from different projections and included these in compiled synthetic
327 datasets.

328 We then contrasted monomer binding to a cluster binding model, which assumed that
329 a whole cluster comprising 7 GFP labeled Mig1 molecules binds a single Mig1 target. Here
330 we randomly assigned the 190 Mig1 molecules into just 27 (i.e. ~190/7) 7-mer clusters to the
331 set of 112 Mig1 target promoter sites. We also implemented improvements of both monomer
332 and cluster binding models to account for the presence of trans-nuclear tracks. Extrapolating
333 the number of detected trans-nuclear foci in our microscope's depth of field over the whole
334 nuclear surface area indicated a total of ~130 Mig1 molecules at *glucose* (+) inside the
335 nucleus prior to export across the cytoplasm. We simulated the presence of these trans-
336 nuclear molecules either using 130 GFP-labeled Mig1 molecules as monomers, or as 18 (i.e.
337 ~130/7) 7-mer clusters at random 3D coordinates over the nuclear envelope surface (Figure. 4
338 – Figure Supplement 3).

339 We discovered that a cluster binding model which included the presence of trans-
340 nuclear foci generated excellent agreement to the experimental foci stoichiometry distribution
341 ($R^2=0.75$) compared to a very poor fit for a monomer binding model ($R^2<0$) (Figure. 4D). The
342 optimized cluster model fit involved on average ~25% of promoter loci to be bound across a
343 population of simulated cells by a 7-mer cluster with the remaining clusters located non-
344 specifically, near the nuclear envelope, consistent with nuclear transit. This structural model
345 supports the hypothesis that the functional unit of Mig1-mediated gene regulation is a cluster
346 of Mig1 molecules, as opposed to Mig1 acting as a monomer.

347

348 **The activator Msn2 also forms functional clusters**

349 We wondered if the discovery of transcription factor clusters was unique to specific
350 properties of the Mig1 repressor, as opposed to being a more general feature of other Zn
351 finger transcription factors. To address this question we prepared a genomically encoded GFP
352 fusion construct of a similar protein Msn2. Nrd1-mCherry was again used as a nuclear
353 marker (Figure. 1 – Figure Supplement 1). Msn2 acts as an activator and not a repressor,
354 which co-regulates several Mig1 target genes but with the opposite nuclear localization
355 response to glucose (18). On performing Slimfield under identical conditions to the Mig1-
356 GFP strain we again observed a significant population of fluorescent Msn2 foci, which had
357 comparable *D* and stoichiometry to those estimated earlier for Mig1 (Table 2). The key

358 difference with the data from the Mig1-GFP strain was that Msn2, unlike Mig1, demonstrated
359 high apparent foci stoichiometry values and lower values of D at *glucose* (-), which was
360 consistent with its role as an activator of the same target genes as opposed to a repressor
361 (Figure. 5A and B). Immuno-gold electron microscopy of fixed Msn2-GFP cells confirmed
362 the presence of GFP in 90nm cryosections with evidence for clusters of comparable
363 diameters to Mig1-GFP (Figure. 4 – Figure Supplement 1C), but with the same technical
364 caveats and poor statistics as reported for the Mig1-GFP dataset. These results support the
365 hypothesis that two different eukaryotic transcription factors that have antagonist effects on
366 the same target genes operate as molecular clusters.

367 To test the functional relevance of Mig1 and Msn2 clusters we performed Slimfield
368 on a strain in which Mig1 and Msn2 were genomically labeled using mCherry and orange
369 fluorescent protein mKO2, respectively (18). This strain also contained a plasmid with GFP
370 labeled PP7 protein to report on nuclear mRNA expressed specifically from the glycogen
371 synthase *GSY1* gene, whose expression can be induced by glucose starvation and is a target of
372 Mig1 and Msn2, labelled with 24 repeats of the PP7 binding sequence (39). In switching
373 from *glucose* (+) to (-) and observing the same cell throughout, we measured PP7
374 accumulating with similar localization patterns to those of Mig1 clusters at *glucose* (+)
375 (Figure. 5C). No accumulation was observed with the mutant Mig1 lacking the Zn finger, in
376 line with previous observations (18). We calculated the numerical overlap integral between
377 these Mig1 and PP7 foci (Figure. 5D), indicating a high mean of ~0.95, where 1 is the
378 theoretical maximum for 100% colocalization in the absence of noise (26). We also observed
379 similar high colocalization between Msn2-mKO2 clusters and PP7-GFP at *glucose* (-)
380 (Figure. 5E). These results demonstrate a functional link between the localization of Mig1
381 and Msn2 clusters, and the transcribed mRNA from their target genes.

382 **Mig1 and Msn2 possess intrinsic disorder which may favor clustering**

383 Since both Mig1 and Msn2 demonstrate significant populations of clustered molecules in
384 functional cell strains we asked the question if there were features common to the sequences
385 of both proteins which might explain this behavior. To address this question we used multiple
386 sequence alignment to determine conserved structural features of both proteins, and
387 secondary structure prediction tools with disorder prediction algorithms. As expected,
388 sequence alignment indicated the presence of the Zn finger motif in both proteins, with
389 secondary structure predictions suggesting relatively elongated structures (Figure. 6A).
390 However, disorder predictions indicated multiple extended intrinsically disordered regions in
391 both Mig1 and Msn2 sequences with an overall proportion of disordered content >50%, as
392 high as 75% for Mig1 (Figure. 6B; Table 6). We measured a trend from a more structured
393 region of Mig1 towards the N-terminus and more disordered regions towards the C-terminus.
394 Msn2 demonstrated a similar bipolar trend but with the structured Zn finger motif towards the
395 C-terminus and the disordered sequences towards the N-terminus. We then ran the same
396 analysis as a comparison against the prokaryotic transcription factor LacI, which represses
397 expression from genes of the *lac* operon as part of the prokaryotic glucose sensing pathway.
398 The predicted disorder content in the case of LacI was <50%. In addition, further sequence
399 alignment analysis predicted that at least 50% of candidate phosphorylation sites in either
400 Mig1 or Msn2 lie within these intrinsically disordered sequences (Table 6; Figure. 6A). An
401 important observation reported previously is that the comparatively highly structured LacI
402 exhibits no obvious clustering behavior from similar high-speed fluorescence microscopy
403 tracking on live bacteria (4). Intrinsically disordered proteins are known to undergo phase
404 transitions which may enable cluster formation and increase the likelihood of binding to
405 nucleic acids (40, 41). It has been shown that homo-oligomerization is energetically more

406 favorable than hetero-oligomerization (42). Moreover, symmetrical arrangement of the same
407 protein can increase accessibility of the protein to binding partners, generate new binding
408 sites, or increase complex specificity and diversity in general (43). We measured significant
409 changes in circular dichroism of the Mig1 fusion construct upon addition of PEG in the
410 wavelength range 200-230nm (Figure. 6C) known to be sensitive to transitions between
411 ordered and intrinsically disordered states (44, 45). Since the Zn finger motif lies towards the
412 opposite terminus to the disordered content for both Mig1 and Msn2 this may suggest a
413 molecular bipolarity which could stabilize a cluster core while exposing Zn fingers on the
414 surface enabling interaction with accessible DNA. This structural mechanism has analogies to
415 that of phospholipid interactions driving micelle formation, however mediated here through
416 disordered sequence interactions as opposed to hydrophobic forces (Figure. 6C). The
417 prevalence of phosphorylation sites located in disordered regions may also suggest a role in
418 mediating affinity to target genes, similar to protein-protein binding by phosphorylation and
419 intrinsic disorder coupling (46).

420

421 **Discussion**

422 Our findings address a totally underexplored and novel aspect of gene regulation with
423 technology that has not been available until recently. In summary, we observe that the
424 repressor protein Mig1 forms clusters which, upon extracellular glucose detection, localize
425 dynamically from the cytoplasm to bind to locations consistent with promoter sequences of
426 its target genes. Similar localization events were observed for the activator Msn2 under
427 glucose limiting conditions. Moreover, Mig1 and Msn2 oligomers colocalized with mRNA
428 transcribed from *GSY1* gene at glucose (+/-), respectively. Our results therefore strongly
429 support a functional link between Mig1 and Msn2 transcription factor clusters and target gene
430 expression. The physiological role of multivalent transcription factor clusters has been
431 elucidated through simulations (9) but unobserved until now. These simulations show that
432 intersegmental transfer between sections of nuclear DNA was essential for factors to find
433 their binding sites within physiologically relevant timescales and requires multivalency.
434 Previous single-molecule studies of p53(47) and TetR(48) in human cancer cells have also
435 suggested a role for non-specific (i.e. sequence independent) transcription factor searching
436 along the DNA. Our findings address the longstanding question of how transcription factors
437 find their targets in the genome so efficiently. Evidence for higher molecular weight Mig1
438 states from biochemical studies has been suggested previously (49). A Mig1-His-HA
439 construct was overexpressed in yeast and cell extracts run in different glucose concentrations
440 through sucrose density centrifugation. In western blots, a higher molecular weight band was
441 observed, attributed to a hypothetical cofactor protein. However, no cofactor was detected
442 and none reported to date. The modal molecular weight observed was ~four times that of
443 Mig1 but with a wide observed distribution consistent with our mean detected cluster size of
444 ~7 molecules. The authors only reported detecting higher molecular weight states in the
445 nucleus in repressing conditions.

446 Clustering of nuclear factors has been reported previously in other systems using
447 single-molecule techniques. In particular, RNA polymerase clustering in the nucleus has been
448 shown to have a functional role in gene regulation through putative transcription factories
449 (50, 51). Other nuclear protein clusters have been shown to have a functional role (52) and
450 the Bicoid transcription factor in *Drosophila melanogaster* embryos has been shown to form
451 clusters partially mediated by regions of intrinsic disorder (53).

452 Our measured turnover of genome-bound Mig1 has similar timescales to that
453 estimated for nucleoid-bound LacI (4), but similar rates of turnover have also been observed
454 in yeast for a DNA-bound activator (54). Faster off rates have been observed during single

455 particle tracking of the DNA-bound fraction of the glucocorticoid receptor (GR) transcription
456 factor in mammalian cells, equivalent to a residence time on DNA of just 1s (12). Single GR
457 molecules appear to bind as a homodimer complex on DNA, and slower Mig1 off rates may
458 suggest higher order multivalency, consistent with Mig1 clusters.

459 Estimating nearest-neighbor distances between Mig1 promoter sites in the *S.*
460 *cerevisiae* genome from the 3C model (Figure. 6D) indicates 20-30% are <50 nm, small
461 enough to enable different DNA segments to be linked through intersegment transfer by a
462 single cluster (6, 9). This separation would also enable simultaneous binding of >1 target
463 (Figure. 6E). The proportion of loci separated by <50nm is also consistent with the estimated
464 proportion of immobile foci and with the proportion of cluster-occupied sites predicted from
465 our structural model. Such multivalency chimes with the tetrameric binding of prokaryotic
466 LacI leading to similar low promoter off rates (4).

467 Measuring the variation of electrostatic charge of residues for the amino acid
468 sequences of both Mig1 and Msn2 (Figure. 6F) we see that the regions in the vicinity of the
469 Zn finger motifs for both proteins have a strong net positive charge compared to the rest of
470 the molecule. If these regions project outwards from a multivalent transcription factor cluster,
471 as per our hypothesized cluster model (Figure. 6E), then the cluster surface could interact
472 electrostatically with the negatively charged phosphate backbone of DNA to enable a 1D
473 sliding diffusion of the protein along a DNA strand, such that the on rate for the protein-DNA
474 interaction is largely sequence-independent in regards to the DNA. Particular details of this
475 type of transcription factor binding to non-specific regions of DNA have been investigated at
476 the level of single transcription factor molecules using computational simulations (55), and
477 suggest initial recognition is most likely via the DNA minor grooves where the phosphates
478 are closer to each other, followed by subsequent interactions between exposed residues on the
479 transcription factor surface and nitrogen bases. This lack of sequence dependence for binding
480 is consistent with observations from an earlier live cell single-molecule tracking study of the
481 TetR repressor (48). We also see experimental evidence for this in our study here, in that we
482 find that the best fit model to account for fluorescence images of the nucleus under high
483 glucose conditions is a combination of occupancy of clusters at the target genes (i.e. sequence
484 specific) with random occupancy to other parts of the genome away from the target genes
485 (i.e. sequence non-specific). Ultimate binding to the gene target once encountered could then
486 be mediated through sequence-specific interactions via the Zn finger motif itself.

487 If the haploid genome of budding yeast, containing 12.1Mbp, is modeled as a flexible
488 ‘virtual’ tube of length 4.1mm ($12.1 \times 10^6 \times 0.34\text{nm}$ for each bp separation parallel to the
489 double helix axis of DNA) with a circular cross-section, then we can calculate that the
490 diameter of the tube required in principle to completely occupy the volume of a typical yeast
491 nucleus (roughly a sphere of diameter $\sim 2\mu\text{m}$) is 30-40nm. This tube diameter, in the absence
492 of local contributions from histone packing, is thus a rough estimate for the effective average
493 separation of DNA strands in the nucleus (i.e. the ‘mesh size’), which is very close to the
494 diameter of clusters we observe. A multivalent transcription factor cluster thus may have only
495 a relatively short distance to diffuse through the nucleoplasm if it dissociates from one DNA
496 strand and then rebinds electrostatically to the next nearest strand, thereby facilitating
497 intersegmental transfer. In this scheme, the association interaction between clusters and
498 neighboring DNA strands is predominantly electrostatic and therefore largely, one might
499 speculate, sequence-independent. However, sequence specificity may be relevant in
500 generating higher-order packed structures of chromatin resulting in localized differences to
501 the nearest neighbor separation of different DNA strands, which could therefore influence the
502 rate at which a cluster transfers from one strand to another. In addition, there may also be
503 localized effects of DNA topology that affect transcription factor binding, which in turn

504 would be expected to have some sequence specificity (55). Also, the off rates of cluster
505 interactions with DNA may be more dependent on the specific sequence. For example, one
506 might anticipate that the dissociation of translocating clusters would be influenced by the
507 presence of obstacles, such as other proteins, already bound to DNA which in turn may have
508 sequence specificity. In particular, bound RNA polymerases present during gene transcription
509 at sequence specific sites could act as roadblocks to kick off translocating clusters from a
510 DNA strand, to again facilitate intersegmental transfer.

511 Several previous experimental studies report observations consistent with
512 intersegmental transfer relevant to our study here. For example, an investigation using single-
513 molecule tracking indicated that transcription factor search times were increased if
514 intersegmental transfer was specifically abrogated (56). These observations are consistent
515 with other experiments that selectively enabled intersegmental transfer by altering DNA
516 conformation (57, 58). Also, they are consistent with biochemical measurements that
517 transcription factors spend a high fraction of their time bound to DNA, as opposed to being in
518 solution (56, 59). Furthermore, other light microscopy studies report direct experimental
519 evidence for intersegmental transfer (6, 60).

520 It is well-established from multiple studies that 3D diffusion of transcription factors in
521 the nucleoplasm alone cannot account for the relatively rapid search times observed
522 experimentally to find specific targets in the genome (3–6). Constraining the dimensionality
523 of diffusion to just 1D, as in the sliding of weakly bound transcription factors on DNA,
524 speeds up this process, but is limited by encountering obstacles already bound to the DNA
525 which potentially result in dissociation of the transcription factor and then slow 3D diffusion
526 in the nucleoplasm. In our system, we speculate that the clusters we observe can slide on
527 DNA in a largely sequence-independent manner but then can cross to neighboring DNA
528 strands in a process likely to have some sequence dependence when an obstacle is
529 encountered, and thus predominantly circumvent the requirement for slow 3D diffusion in the
530 nucleoplasm. Minimizing the contribution from the slowest component in the search process
531 may therefore result in an overall reduction in the amount of time required for a given
532 transcription factor to find its gene target.

533 Extensive bioinformatics analysis of proteome disorder across a range of species
534 suggests a sharp increase from prokaryotes to eukaryotes (61), speculatively due to the
535 prokaryotic absence of cell compartments and regulated ubiquitination mechanisms lowering
536 protection of unfolded disordered structures from degradation (62). Our discovery in yeast
537 may reveal a eukaryotic adaptation that stabilizes gene expression. The slow off rate we
538 measure would result in insensitivity to high frequency stochastic noise which could
539 otherwise result in false positive detection and an associated wasteful expression response.
540 We also note that long turnover times may facilitate modulation between co-regulatory
541 factors by maximizing overlap periods, as suggested previously for Mig1/Msn2 (18).

542 Our results suggest that cellular depletion forces due to crowding enable cluster
543 formation. Crowding is known to increase oligomerization reaction rates for low association
544 proteins but slow down fast reactions due to an associated decrease in diffusion rates, and
545 have a more pronounced effect on higher order multimers rather than dimers (36). It is
546 technically challenging to study depletion forces *in vivo*, however there is growing *in vitro*
547 and *in silico* evidence of the importance of molecular crowding in cell biology. A particularly
548 striking effect was observed previously in the formation of clusters of the bacterial cell
549 division protein FtsZ in the presence of two crowding proteins – hemoglobin and BSA (63).
550 Higher order decamers and multimers were observed in the presence of crowding agents and
551 these structures are thought to account for as much as 1/3 of the *in vivo* FtsZ content.

552 Similarly, two recent yeast studies of the high-osmolarity glycerol (HOG) pathway also
553 suggest a dependence on gene expression mediated by molecular crowding (64, 65).

554 The range of GFP labeled Mig1 cluster diameters *in vivo* of 15-50nm is smaller than the
555 80nm diameter of yeast nuclear pore complexes (66), not prohibitively large as to prevent
556 intact clusters from translocating across the nuclear envelope. An earlier *in vitro* study using
557 sucrose gradient centrifugation suggested a Stokes radius of 4.8 nm for the Mig1 fraction, i.e.
558 diameter 9.6nm, large for a Mig1 monomer (49) whose molecular weight is 55.5kDa, e.g. that
559 of monomeric bovine serum albumin (BSA) at a molecular weight of 66kDa is closer to
560 3.5nm (67). The authors ascribed this effect to a hypothetical elongated monomeric structure
561 for Mig1. The equivalent Stokes radius for GFP has been measured at 2.4nm (68), i.e.
562 diameter 4.8nm. Also, for our Mig1-GFP construct there are two amino acids residues in the
563 linker region between the Mig1 and GFP sequences (i.e. additional length 0.7-0.8nm). Thus
564 the anticipated hydrodynamic diameter of Mig1-GFP is 15-16nm. The mean observed ~7-mer
565 cluster diameter from Slimfield data is ~30nm, which, assuming a spherical packing
566 geometry, suggests a subunit diameter for single Mig1-GFP molecules of $\sim 30/7^{1/3} \approx 15.6$ nm,
567 consistent with that predicted from the earlier hydrodynamic expectations. Using Stokes law
568 this estimated hydrodynamic radius indicates an effective viscosity for the cytoplasm and
569 nucleoplasm as low as 2-3cP, compatible with earlier live cell estimates on mammalian cells
570 using fluorescence correlation spectroscopy (FCS) (69).

571 One alternative hypothesis to that of intrinsically disordered sequences mediating Mig1
572 cluster formation is the existence of a hypothetical cofactor protein to Mig1. However, such a
573 cofactor would be invisible on our Slimfield assay but would result in a larger measured
574 hydrodynamic radius than we estimate from fluorescence imaging, which would be manifest
575 as larger apparent viscosity values than those we observe. Coupled to observations of Msn2
576 forming clusters also, and the lack of any reported stable cofactor candidate to date, limits the
577 cofactor hypothesis. Pull down assays do suggest that promoter bound Mig1 consists of a
578 complex which includes the accessory proteins Ssn6 and Tup1 (35), however this would not
579 explain the observation of Mig1 clusters outside the nucleus.

580 There may be other advantages in having a different strategy between *S. cerevisiae* and
581 *E. coli* to achieve lowered transcriptional regulator off rate. A clue to these may lie in
582 phosphorylation. We discovered that at least 50% of candidate serine or threonine
583 phosphorylation sites in Mig1 and Msn2 lie in regions with high intrinsic disorder, which
584 may have higher sequence-unspecific binding affinities to DNA (40, 41). Thus
585 phosphorylation at sites within these regions may potentially disrupt binding to DNA, similar
586 to observed changes to protein-protein affinity being coupled to protein phosphorylation state
587 (46). Previous studies indicate that dephosphorylated Mig1 binds to its targets (70). Thus,
588 intrinsic disorder may be required for bistability in affinity of Mig1/Msn2 to DNA.

589 Wide scale bioinformatics screening reveals a significant prevalence of intrinsic disorder
590 in eukaryotic transcription factors (71). Our discovery is the first, to our knowledge, to make
591 a link between predicted disorder and the ability to form higher-order clusters in transcription
592 factors. Thus, our results address the longstanding question of why there is so much predicted
593 disorder in eukaryote transcription factors. Our observations that protein interactions based
594 on weak intracellular forces and molecular crowding has direct functional relevance may
595 stimulate new research lines in several areas of cell biology. For example, our findings may
596 have important mechanistic implications for other aggregation processes mediated through
597 intrinsic disorder interactions, such as those of amyloid plaques found in neurodegenerative
598 disorders including Alzheimer's and Parkinson's diseases (72). Increased understanding of
599 the clustering mechanism may not only be of value in understanding such diseases, but could
600 enable future novel synthetic biology applications to manufacture gene circuits with, for
601 example, a range of bespoke response times.

602

603 **Materials and methods**

604

605 **Strain construction and characterization**

606 We developed Mig1 fluorescent protein strains based on strain YSH1351 (16) using eGFP in
607 the first instance and also mGFP/GFPmut3 designed to inhibit oligomerization (23), and
608 photoswitchable mEos2 (24). Mig1-mGFP and Mig1-mEos2 fusions were constructed by
609 introducing into YSH1351 (BY4741 wild type) cells the *mGFP-HIS3* or *mEos2-HIS3* PCR
610 fragment flanked on its 5' end with 50bp sequence of *MIG1* 3' end and 50bp downstream of
611 *MIG1* excluding the STOP codon. The *mEos2-HIS3* and *mGFP-HIS3* fragment was
612 amplified from mEOS-his plasmid (GeneArt, Life Technologies) and pmGFP-S plasmid
613 designed for this study by inserting the mGFP sequence into plasmid YDp-H. Modified
614 strains in which the *SNF1* gene was deleted, *snf1Δ*, were prepared by compromising the gene
615 with an auxotrophic marker by providing the *LEU2* fragment amplified from plasmid YDp-L
616 and flanked with 50bp of *SNF1* upstream and downstream sequence on 5' and 3' ends,
617 respectively, directly into cells. Strains in which Snf1 kinase activity can be inhibited by
618 25μM 1NM-PP1 in DMSO were prepared by introducing into cells a plasmid with an ATP
619 analog-sensitive version of Snf1 with *I132G* mutation (73). DMSO itself has been shown
620 previously not to affect Mig1's behavior under different glucose conditions (17) similar to
621 our own findings (Figure. 2 – Figure supplement 2). All transformations were performed
622 using the lithium acetate protocol (74).

623 Cell doubling times of all strains were calculated (75) (Figure. 1 – Figure Supplement
624 2A) based on OD₆₀₀ values obtained during cultivation in media supplemented with 4% or
625 0.2% glucose (Bioscreen analyser C). We quantified mRNA relative expression of the
626 *MIG1* gene using qPCR against the constitutive actin gene *ACT1* in the wild type and the
627 Mig1-mGFP strain in cells pre-grown in 4% glucose and then shifted to elevated (4%)
628 and depleted (0.2%) extracellular glucose for 2 h. mRNA isolation and cDNA synthesis were
629 performed as described previously (76).

630 For Msn2-GFP experiments we used the YSH2350 strain (*MATa msn2-GFP-HIS3 nrd1-*
631 *mCherry-hphNT1 MET LYS*) in BY4741 background.

632

633 **Protein production and purification**

634 His-tagged *mCherry*, *eGFP* and *mGFP* genes were amplified by PCR and cloned into pET
635 vectors. An expression pRSET A plasmid containing 6xHis-Mig1-mGFP was obtained
636 commercially (GeneArt, Life Technologies). *Escherichia coli* strain BL21(DE3) carrying the
637 expression plasmid was grown in LB with 100μg/ml ampicillin and 34μg/ml
638 chloramphenicol at 37°C to OD₆₀₀ 0.7. Protein expression was induced by adding isopropyl-
639 β-D-thiogalactopyranoside (IPTG) at final concentration of 1mM for 3h at 30°C. Cells were
640 suspended in 50mM NaH₂PO₄, 10mM Tris, 300mM NaCl, 2mM EDTA, 0.2mM PMSF,
641 0.1% β-mercaptoethanol, pH 8.0, and lysed by sonication or by three passages through a
642 chilled Emulsiflex (Avestin). Extracts were cleared (24,000g, 30min) and filtered (pore
643 diameter 0.45μm; Millipore, Bedford). All proteins were purified using Ni²⁺ affinity
644 chromatography on a 5ml HisTrap FF column (GE Healthcare). Mig1-mGFP was eluted with
645 a linear gradient 0 - 0.4 M imidazole in lysis buffer. Mig1-mGFP was further purified by
646 size-exclusion chromatography (Superdex 200 Increase 10/300, GE Healthcare) and
647 concentrated (50 kDa molecular weight cutoff VIVASPIN 20 concentrator). Purity of the
648 sample was confirmed by Coomassie stained SDS-PAGE gels (Simply Blue Safe Stain, Life
649 Technologies).

650

651 **Media and growth conditions**

652 Cells from frozen stocks were grown on plates with standard YPD media (10 g/l yeast
653 extract, 20 g/l bacto-peptone, 20 g/l agar) supplemented with 4% glucose (w/v) at 30⁰C
654 overnight. For the liquid cultures, the YPD was prepared as above but without agar, and the
655 cells were grown at 30⁰C while shaking (180 rpm).

656 For transformants that carried a plasmid with mutated *SNF1* (p*SNF1-I132G*) or PP7-
657 2xGFP (pDZ276), minimal YNB media with –URA amino acid supplement was applied. For
658 the growth rate experiments cells were grown on 100 well plates in YNB with complete
659 amino acid supplement and 4% glucose (w/v) until logarithmic phase, subcultured into fresh
660 medium on a new 100 well plate and grown until logarithmic phase again. 10 µl of each
661 culture was resuspended in 250 µl of fresh YNB medium with 4% or 0.2% glucose (w/v) on a
662 new plate and cultivated in Bioscreen analyser C for 96 h at 30⁰C or 22⁰C. OD measurements
663 at 600 nm were taken every 10 min with prior shaking. Each strain was represented in
664 sextuplicates.

665 For microscopy experiments on the BY4741 wild type and/or cells with genetically
666 integrated fluorescent proteins, minimal YNB media (1.7 g/l yeast nitrogen base without
667 amino acids and (NH₄)₂SO₄, 5 g/l (NH₄)₂SO₄, 0.79 g/l complete amino acid supplement as
668 indicated by manufacturer) with appropriate glucose concentrations was used. In brief, cells
669 were first streaked onto YPD plates, grown overnight at 30⁰C prior to culturing in liquid
670 minimal YNB media with complete amino acid supplement and 4% glucose overnight, then
671 sub-cultured into fresh YNB with 4% glucose for 4h with shaking at 30⁰C. Cultures were
672 spun at 3,000rpm, re-suspended into fresh YNB with (4%) or without (0%) glucose,
673 immobilized in 1µl spots onto an 1% agarose well perfused with YNB minimal media with
674 an appropriate glucose concentration enclosed between a plasma-cleaned BK7 glass
675 microscope coverslip and slide, which permitted cells to continue to grow and divide (20, 21)
676 while being observed for up to several hours if required. Images were acquired not longer
677 than 2 hours after the last media switch.

678

679 **SDS-PAGE**

680 50 ml cultures of YSH1703 transformed with centromeric pMig1-HA and pSNF1-I132G-
681 TAP or pSNF1-TAP plasmids were grown until mid-log phase in yeast nitrogen base, 4%
682 glucose, uracil and histidine deficient. Each culture was separated into two new cultures with
683 4% and 0.05% glucose, respectively, and incubated for 30 min. The following procedure was
684 adapted from Bendrioua et al (16). Cells were harvested by centrifugation (3,000rpm, 50s),
685 suspended in 1 ml of 0.1M NaOH for 5 min and spun down. Pellets were suspended in 2 ml
686 of 2M NaOH with 7% β- mercaptoethanol for 2 min and then 50% trichloroacetic acid was
687 added. Samples were vortexed and spun down at 13,000rpm. The pellets were washed in
688 0.5 ml of 1M Tris-HCl (pH 8.0), resuspended in 50 µl of 1x SDS sample buffer (62.5 mM
689 Tris-HCl (pH 6.8), 3% SDS, 10% glycerol, 5% β-mercaptoethanol, and 0.004% bromophenol
690 blue) and boiled for 5 min. The protein extracts were obtained by centrifuging at the maximal
691 speed and collecting the supernatants. For western blotting, 50 µg of extracted proteins were
692 resolved on a Criterion TGX 10% precast polyacrylamide gel, then transferred onto a
693 nitrocellulose membrane (Trans-Blot Turbo Transfer Pack, Bio-Rad) using Trans-Blot Turbo
694 Transfer System (Bio-Rad). After transfer, the membrane was blocked in Odyssey Blocking
695 buffer (LI-COR Biosciences). Mig1 was detected using primary mouse anti-HA (1:2000)
696 antibodies (Santa Cruz), then secondary goat anti-mouse IRDye-800CW (1:5000) antibodies
697 (LI-COR Biosciences). The result was visualized using an infrared imager (Odyssey, LI-COR
698 Biosciences), 800nm channel.

699

700

701 **Native PAGE**

702 A 50 ml culture of the YSH2862 strain was grown until mid-log phase in rich media with 4%
703 glucose, then, 25 ml of the culture was transferred into fresh YPD with 4% glucose, and the
704 rest into YPD with 0.05% glucose for 30 min. The cultures were harvested by centrifugation,
705 suspended in 0.1ml of solubilization buffer (100 mM Tris-HCl, pH 6.8, 0.1 mM Na₃VO₄, 1x
706 protease inhibitor cocktail (Roche), 0.1% Triton-X100). 400µl of glass beads were added, and
707 cells were broken by FastPrep, 6m/s, 20 s. Protein extracts were obtained by adding 150 µl of
708 solubilization buffer, centrifugation at 13,000 rpm, 5min and collecting the supernatant.
709 Protein quantification was performed by using Bradford with BSA standard (Bio-Rad). 250
710 µg of total protein extracts were run on a Criterion TGX Stain Free 10% precast
711 polyacrylamide gel (Bio-Rad). Samples were diluted 1:1 with 2x Native Sample Buffer (Bio-
712 Rad). Electrophoresis was performed at 4⁰C starting at 100V until the bromophenol blue line
713 reached the end of the gel. The gel was transferred onto a nitrocellulose membrane (Trans-
714 Blot Turbo Transfer Pack, Bio-Rad) using Trans-Blot Turbo Transfer System (Bio-Rad).
715 After transfer, the membrane was blocked in Odyssey Blocking buffer (LI-COR
716 Biosciences), analyzed by immunoblotting with mouse anti-GFP (1:500) antibodies (Roche)
717 and visualized with goat anti-mouse IRDye-800CW (1:5,000) antibodies (LI-COR
718 Biosciences) by using an infrared imager (Odyssey, LI-COR Biosciences), 800nm channel.
719 As a molecular weight reference, a NativeMark Unstained Protein Standards (Invitrogen)
720 were used.

721

722 **Slimfield microscopy**

723 A dual-color bespoke laser excitation single-molecule fluorescence microscope was used (21,
724 27, 77) utilizing narrow epifluorescence excitation of 10µm full width at half maximum
725 (FWHM) in the sample plane to generate Slimfield illumination. GFP and mCherry excitation
726 used co-aligned linearly polarized 488 nm and 561 nm wavelength 50 mW lasers (Coherent
727 Obis) respectively which could be attenuated independently via neutral density filters
728 followed by propagation through an achromatic $\lambda/2$ plate to rotate the plane of polarization
729 prior to separation into two independent paths generated by splitting into orthogonal
730 polarization components by a polarization splitting cube to enable simultaneous Slimfield
731 illumination and a focused laser bleach illumination path for fluorescence recovery after
732 photobleaching (FRAP) when required. The two paths were reformed into a single common
733 path via a second polarization cube, circularized for polarization via an achromatic $\lambda/4$ plate
734 with fast axis orientated at 45° to the polarization axes of each path and directed at ~6 W/cm²
735 excitation intensity onto the sample mounted on an xyz nanostage (Mad City Labs) via a dual-
736 pass green/red dichroic mirror centered at long-pass wavelength 560nm and emission filters
737 with 25nm bandwidths centered at 525nm and 594nm (Chroma).

738 Fluorescence emissions were captured by a 1.49NA oil immersion objective lens
739 (Nikon) and split into green and red detection channels using a bespoke color splitter utilizing
740 a long-pass dichroic mirror with wavelength cut-off of 565nm prior to imaging each channel
741 onto separate halves of the same EMCCD camera detector (iXon DV860-BI, Andor
742 Technology, UK) at a pixel magnification of 80 nm/pixel using 5ms camera exposure time.
743 We confirmed negligible measured crosstalk between GFP and mCherry signals to red and
744 green channels respectively, using purified GFP and mCherry sampled in an *in vitro* surface
745 immobilization assay (details below).

746 Three color microscopy was performed on the same microscope, using a 50mW 532nm
747 wavelength laser (Obis) to excite mKO2, coupled into the same optics as before with the
748 addition of a 532nm notch rejection filter (Semrock) in both channels of the imaging path.
749 This allowed 1mW of laser excitation at the sample. Due to the high copy number of plasmid
750 expressed PP7-2xGFP and the 48 RNA loci, the 488nm wavelength laser was attenuated to

751 ~10 μ W. Each fluorophore was separately excited in the following order: mCherry, mKO2
752 and GFP to prevent crosstalk. mCherry and mKO2 both emit in the ‘red’ channel of the
753 microscope, while GFP appears in the ‘green’ with very limited crosstalk.

754

755 **Microfluidics control of single cell imaging**

756 To investigate time-resolved glucose concentration-dependent changes in Mig1-GFP
757 localization in individual yeast cells, we used bespoke microfluidics and our bespoke control
758 software *CellBild* (LabVIEW, National Instruments), enabling cell-to-cell imaging in
759 response to environmental glucose changes. *CellBild* controlled camera acquisition
760 synchronized to flow-cell environmental switches via a syringe pump containing an alternate
761 glucose environment. Microfluidic flow-chambers were based on an earlier 4-channel design
762 (78).

763 Prior to each experiment flow-chambers were wetted and pre-treated for 15min with
764 1 mg/ml of concanavalin A (ConA) which binds to the glass surface of the plasma cleaned
765 flow-chamber. Cells were introduced via a side channel and were left to bind ConA for
766 15min to immobilize cells on the surface. Any remaining ConA and unbound cells were
767 washed out and a steady flow of YNB with 0% glucose provided to one of the central
768 channels by gravity feed. A syringe pump synchronized with image acquisition introduced
769 YNB with 4% glucose in the second central channel. The pumped alternate environment
770 reaches cells within 1-2s at a flow rate of 10 μ l/min, enabling rapid change between two
771 different glucose concentrations.

772 Slimfield imaging was performed on a similar bespoke microscope setup at comparable
773 laser excitation intensities and spectral filtering prior to imaging onto a Photometrics *Evolve*
774 *Delta 512* EMCCD camera at 200 frames per second. Alternating frame laser excitation
775 (ALEX) was used to minimize any autofluorescence contamination in the red channel
776 introduced by the blue excitation light.

777 Around 1-4 cells were imaged in a single field of view for each glucose exchange. The
778 same flow chamber was used for multiple fields of view such that each cell analyzed may
779 have experienced up to four glucose exchange cycles.

780

781 **Foci detection, tracking and stoichiometry determination**

782 Foci were automatically detected using software written in MATLAB (Mathworks) (25),
783 lateral localization ~40nm, enabling estimates of D and stoichiometry. Our bespoke foci
784 detection and tracking software objectively identifies candidate bright foci by a combination
785 of pixel intensity thresholding and image transformation to yield bright pixel coordinates.
786 The intensity centroid and characteristic intensity, defined as the sum of the pixel intensities
787 inside a 5 pixel radius region of interest around the foci minus the local background and
788 corrected for non-uniformity in the excitation field are determined by iterative Gaussian
789 masking. If the signal-to-noise ratio of the foci, defined as the characteristic intensity per
790 pixel/background standard deviation per pixel, is >0.4 it is accepted and fitted with a 2D
791 radial Gaussian function to determine its sigma width, which our simulations indicate single-
792 molecule sensitivity under typical *in vivo* imaging conditions (27). Foci in consecutive image
793 frames within a single point spread function (PSF) width, and not different in brightness or
794 sigma width by more than a factor of two, are linked into the same track. The microscopic
795 diffusion coefficient D is then estimated for each accepted foci track using mean square
796 displacement analysis, in addition to several other mobility parameters.

797 Cell and nuclear boundaries were segmented from GFP and mCherry fluorescence
798 images respectively using a relative threshold pixel intensity value trained on simulated
799 images of uniform fluorescence in idealized spherical compartments. An optimized threshold

800 value of 0.3 times the mean compartment fluorescence intensity segmented the boundary to
801 within 0.5 pixels.

802 The characteristic brightness of a single GFP molecule was determined directly from *in*
803 *vivo* data and corroborated using *in vitro* immobilized protein assays (22). The intensity of
804 tracked fluorescent foci in live cells was measured over time as described above (Figure. 1–
805 Figure Supplement 3). These followed an approximately exponential photobleach decay
806 function of intensity with respect to time. Every oligomeric Mig1-GFP complex as it
807 photobleaches to zero intensity will emit the characteristic single GFP intensity value, I_{GFP} ,
808 i.e. the brightness of a single GFP molecule, given in our case by the modal value of all foci
809 intensities over time, and can potentially bleach in integer steps of this value at each sampling
810 time point. This value of I_{GFP} was further verified by Fourier spectral analysis of the pairwise
811 distance distribution (22) of all foci intensities which yields the same value to within
812 measurement error in our system.

813 All foci tracks found within 70 image frames of the start of laser illumination were
814 included in the analysis and were corrected for photobleaching by weighting the measured
815 foci intensity I at a time t following the start of laser illumination with a function $\exp(+t/t_b)$ to
816 correct for the exponential photobleach decay $I_0\exp(-t/t_b)$, of each intensity trace with a fixed
817 time constant, where I_0 is the initial unbleached intensity. This photobleach time constant t_b
818 was determined from exponential decay fits to the foci intensities and whole cell intensities
819 over time to be 40 ± 0.6 ms. Stoichiometries were obtained by dividing the photobleach
820 estimate for the initial intensity I_0 of a given foci by the characteristic single GFP molecule
821 brightness value I_{GFP} .

822 Autofluorescence correction was applied to pool quantification by subtracting the red
823 channel image from the green channel image multiplied by a correlation factor. By
824 comparing wild type and GFP cell images we confirmed that when only the GFP exciting
825 488 nm wavelength laser was used the green channel image contained fluorescence intensity
826 from GFP and autofluorescence, while the red channel contains only autofluorescence pixels,
827 consistent with expectations from transmission spectra of known autofluorescent components
828 in yeast cells. We measured the red channel autofluorescence pixels to be linearly
829 proportional to the green channel autofluorescence pixels. The scaling factor between
830 channels was determined by Slimfield imaging of the wild type yeast strain (i.e. non GFP)
831 under the same conditions and comparing intensity values pixel-by-pixel in each channel. A
832 linear relationship between pixels was found with scaling factor of 0.9 ± 0.1 .

833 Copy numbers of Mig1-GFP of the pool component were estimated using a previously
834 developed CoPro algorithm (27). In brief, the cytoplasmic and nuclear pools were modelled
835 as uniform fluorescence over spherical cells and nuclei using experimentally measured radii.
836 A model PSF was integrated over these two volumes to create model nuclear and cytoplasmic
837 images and then their relative contributions to the camera background and autofluorescence
838 corrected GFP intensity image determined by solving a set of linear equations for each pixel.
839 Dividing the contributions by the characteristic single GFP molecule intensity and correcting
840 for out-of-plane foci yields the pool concentration.

841 Stoichiometry distributions were rendered as objective kernel density estimations (22)
842 using a Gaussian kernel with bandwidth optimized for normally distributed data using
843 standard MATLAB routines.

844

845 **Stochastic optical reconstruction microscopy (STORM)**

846 To photoswitch Mig1-mEos2, a 405 nm wavelength laser (Coherent Obis), attenuated to
847 $\sim 1\text{mW}/\text{cm}^2$ was used in conjunction with the 488 nm and 561 nm lasers on the Slimfield
848 microscope, similar to previous super-resolution imaging of yeast cells (79). The 405 nm
849 laser light causes mEos2 to photoswitch from a green (excitable via the 488 nm laser) to a red

850 (excitable by the 561 nm laser) fluorescent state. Using low intensity 405 nm light generates
851 photoactive fluorophore foci, photobleached by the 561 nm laser at a rate which results in an
852 approximately steady-state concentration density in each live cell studied. The bright foci
853 were tracked using the same parameters and criteria for spot acceptance as the Slimfield data.
854 The tracks were then used to generate a super-resolved image heat map with 20nm pixel size
855 by the summation of 2D Gaussian functions at each sub-pixel. Here, we assumed a sigma
856 width of the 2D Gaussian function of 40nm to match the measured lateral precision following
857 automated particle tracking of Mig1-mEos2 foci (27).

858

859 **Fluorescent protein brightness characterization**

860 We used a surface-immobilization assay described previously (21, 27) employing antibody
861 conjugation to immobilize single molecules of GFP respectively onto the surface of plasma-
862 cleaned BK7 glass microscope coverslips and imaged using the same buffer medium and
863 imaging conditions as for live cell Slimfield experiments, resulting in integrated single-
864 molecule peak intensity values for mGFP of $4,600 \pm 3,000$ (\pm half width half maximum,
865 HWHM) counts. Similar experiments on unmodified purified Clontech eGFP generated peak
866 intensity values of $4,700 \pm 2,000$ counts, statistically identical to that of mGFP (Student *t*-test,
867 $p=0.62$) with no significant indication of multimerization effects from the measured
868 distribution of foci intensity values. Similarly, Slimfield imaging and foci stoichiometry
869 analysis on Mig1-mGFP and Mig1-eGFP cell strains were compared *in vivo* under high and
870 low glucose conditions in two separate cell strains, resulting in distributions which were
871 statistically identical (Pearson's χ^2 test comparing KDEs, Figure. 1 - Figure Supplement 2E
872 and F). These results indicated no measurable differences between multimerization state or
873 single-molecule foci intensity between mGFP and eGFP which enabled direct comparison
874 between Mig1-eGFP cell strain data obtained from preliminary experiments here and from
875 previous studies (16).

876 Maturation effects of mCherry and GFP were investigated by adding mRNA translation
877 inhibitor antibiotic cycloheximide, final concentration 100 $\mu\text{g/ml}$, for 1h (80), photobleaching
878 cells, then monitoring any recovery in fluorescence as a metric for newly matured fluorescent
879 material in the cell. Cells were prepared for microscopy as before but using cycloheximide in
880 all subsequent preparation and imaging media and imaged using a commercial mercury-arc
881 excitation fluorescence microscope Zeiss Axiovert 200M (Carl Zeiss MicroImaging) onto an
882 ApoTome camera using a lower excitation intensity than for Slimfield imaging but a larger
883 field of view, enabling a greater number of cells to be imaged simultaneously.

884 Surface-immobilized cells using strain YSH2863 were photobleached by continuous
885 illumination for between 3min 40s to 4min until dark using separate filter sets 38HE and
886 43HE for GFP and mCherry excitation, respectively. Fluorescence images were acquired at
887 subsequent time intervals up to 120min and analyzed using AxioVision software (Figure. S6).
888 The background-corrected total cellular fluorescence intensity was quantified at each time
889 point for each cell using ImageJ software. Comparison between Mig1-GFP fluorescence
890 signal and the green channel signal from the parental strain BY4741, and the Nrd1-mCherry
891 signal and the red channel signal from the parental strain, indicate fluorescence recovery after
892 correction above the level of any autofluorescence contributions of <15% for GFP and
893 mCherry over the timescale of our experiments, consistent with previous estimates of *in vivo*
894 maturation times for GFP and mCherry (21, 22, 81).

895

896 **Characterizing Mig1-GFP clusters *in vitro***

897 Using Slimfield microscopy under the same imaging conditions as for live cell microscopy
898 we measured the fluorescent foci intensity of 1 $\mu\text{g/ml}$ solutions of purified Mig1-mGFP and
899 mGFP using the normal imaging buffer of PBS, compared with the imaging buffer

900 supplemented with 1kDa molecular weight PEG at a concentration of 10% (w/v) used to
901 reproduce cellular depletion forces (36, 75).

902

903 **Circular dichroism measurements**

904 Purified Mig1-mGFP was placed in 25 mM Na₂HPO₄, pH 7.0, by buffer exchange procedure
905 with a Pur-A-Lyser Maxi dialysis Kit (Sigma Aldrich) for 3h at 4⁰C with constant stirring in
906 500 ml buffer. Circular dichroism measurements were performed on a Jasco J810 circular
907 dichromator with Peltier temperature control and Biologic SFM300 stop-flow accessory on
908 0.16mg/ml Mig1-mGFP samples with or without 20% PEG-1000 at 20⁰C, from 260 to 200
909 nm, a 2 nm band width, 2 sec response time, at the speed of 100 nm/min. The resulting
910 spectrum represents the average of 5 scans, indicating a typical SD error of ~0.1 mdeg
911 ellipticity. Spectra from 25 mM Na₂HPO₄ and 25 mM Na₂HPO₄ with 20% (w/v) PEG were
912 used as a background and subtracted from spectra of Mig1-mGFP without or with 20% (w/v)
913 PEG respectively.

914

915 **Immuno-gold electron microscopy**

916 Cells for Mig1-GFP and Msn2-GFP strains as well as the wild type control strain containing
917 no GFP were grown using the same conditions as for Slimfield imaging but pelleted down at
918 the end of growth and prepared for immuno electron microscopy using an adaptation of the
919 Tokuyasu cryosectioning method (82) following the same protocol that had been previously
920 optimized for budding yeast cells (83) to generate ~90nm thick cryosections, with the
921 exception that the sections were picked up on a drop of 2.3M sucrose, placed on the grid, then
922 floated down on PBS, and then immunolabeled immediately, rather than storing on gelatine
923 as occurred in the earlier protocol. The grids used were nickel, with a formvar/carbon support
924 film. In brief, the immunolabeling protocol used a 0.05M glycine in PBS wash of each
925 section for 5 min followed by a block of 10% goat serum in PBS (GS/PBS) pre-filtered
926 through a 0.2 μm diameter filter. Then an incubation of 1 h with the primary antibody of
927 rabbit polyclonal anti-GFP (ab6556, Abcam) at 1 in 250 dilution from stock in GS/PBS. Then
928 five 3 min washes in GS/PBS. Then incubation for 45 min with the goat anti-IgG-rabbit
929 secondary antibody labeled with 10nm diameter gold (EM.GAR10, BBI solutions) at a
930 dilution of 1 in 10 from stock. Sections were then washed five more times in GS/PBS prior to
931 chemical fixation in 1% glutaraldehyde in sodium phosphate buffer for 10 min, then washed
932 in dH₂O five times for 3 min each and negative-stained using methyl cellulose 2% in 0.4%
933 uranyl acetate, and then washed twice more in dH₂O prior to drying for 10 min. Drop sizes for
934 staining, blocking and washing onto sections were 50 μl, while antibody incubations used
935 25 μl drops, all steps performed at room temperatures.

936 Electron microscopy was performed on these dried sections using a 120kV Tecnai 12
937 BioTWIN (FEI) electron microscope in transmission mode, and imaged onto an SIS
938 Megaview III camera. From a total of ~150 control cells containing no GFP we could detect
939 no obvious signs of gold labeling. Using approximately the same number of cells for each of
940 the Mig1-GFP and Msn2-GFP strains all images showed evidence for at least one gold foci
941 labeling in the cytoplasm, though labeling was largely absent from the nucleus possibly due
942 to poor antibody accessibility into regions of tightly packed DNA since the combined Stokes
943 radii from the primary and secondary antibodies is comparable to the mean effective DNA
944 mesh size in the yeast nucleus of a few tens of nm (see Discussion section). We estimate that
945 the thin cryosections occupy ~2.5% of the volume of an average yeast cell and so based on
946 our copy number estimates from fluorescence microscopy in the accessible cytoplasmic
947 compartment the maximum number of GFP available for labelling in each cryosection is ~20
948 molecules. We observed a range of 1-8 gold foci in total per cell across the GFP datasets and
949 so the overall labelling efficiency in these experiments is low at typically 20% or less.

950 However, we observed 10 cells from a set of ~150 from each of the Mig-GFP and Msn2-GFP
951 strains (i.e. ~7% of the total) which showed >1 gold foci clustering together inside an area of
952 effective diameter ~50nm or less, with up to 7 gold foci per cluster being observed.

953

954 **Bioinformatics analysis and structural modeling**

955 Bioinformatics analysis was used to identify candidate promoter sequences in the budding
956 yeast genome. The Mig1 target pattern sequence was identified based on 14 promoter
957 sequences (37) using the IUPAC nucleotide code. The entire *S. cerevisiae* S288c genome was
958 scanned in order to find all sequences that matched the pattern. The scanning was performed
959 by RNABOB software (84), and collated for any further analysis and identification of the
960 sequences lying within promoter regions. All information regarding *S. cerevisiae* genes was
961 obtained from SGD database (<http://yeastgenome.org/>).

962 Using a consensus structural model for the budding yeast chromosome based on 3C
963 data (38) we explored various different models of Mig1 binding to the putative promoter
964 sequence identified from the bioinformatics analysis. We generated simulated images from
965 these models adding experimentally realistic levels of signal and noise, and ran these data
966 through the same foci detection and analysis software as for the real live cell data using
967 identical parameters throughout. We then compared these results to the measured
968 experimental stoichiometry (Figure. 4C). Monomer models assume that a single Mig1
969 molecule binds to a target promoter site, whereas cluster models assume that a cluster
970 comprising 7 Mig1 molecules (based on our observations of stoichiometry periodicity) binds
971 a single target promoter. Copy number analysis indicated 190 Mig1 molecules per cell on
972 average associated with foci. In the monomer model (Fig. 4C) all 109 promoter sites were
973 assigned a Mig1 molecule and the remaining 81 Mig1 molecules were placed randomly in the
974 222 remaining Mig1 target binding sites within the rest of the genome. In the DNA cluster
975 model (Figure. 4 – Figure Supplement 1) we randomly assigned the observed 190 Mig1
976 molecules in foci into just 27 clusters to Mig1 target promoter sites. We also tested two
977 nuclear envelope (NE) variants of both models, to account for the trans-nuclear tracks: here,
978 typically ~7 Mig1 were observed translocating from the nucleus to the cytoplasm at *glucose*
979 (+) within the microscope's depth of field; extrapolating this value over the whole nucleus
980 this indicates ~130 Mig1 molecules within the nucleus but less than a single PSF width from
981 the nuclear envelope prior to export to the cytoplasm. We simulated this effect using either
982 130 Mig1 molecules as Mig1 monomers or as 18 (i.e. ~130/7) 7-mer clusters at random 3D
983 coordinates at the simulated nuclear envelope position in the 3C model. Finally, to generate
984 the best fit Mig1 cluster model, we obtained an optimized fit to the data with a mixed
985 population model with 75% of cells in the NE cluster model and 25% in the DNA cluster
986 model. We note here that the monomer model can produce higher apparent stoichiometry due
987 to the increased density of resulting foci (although the same density of Mig1).

988 We used bioinformatics to investigate the extent of intrinsic disorder in the amino acid
989 sequence of budding yeast Mig1 and Msn2 proteins as well as the *E. coli lac* repressor LacI,
990 employing the Predictor of Natural Disordered Regions (PONDR) algorithm (85) (online tool
991 <http://www.pondr.com/cgi-bin/PONDR/pondr.cgi>) with a VL-XT algorithm. We also used
992 the secondary structure prediction algorithm of PyMOL
993 (<http://www.pymolwiki.org/index.php/Dss>) to highlight disordered and structured regions and
994 display the unfolded protein chain, and used PSI-BLAST multiple sequence alignment to
995 determine conserved structural features of Mig1 for the Zn finger motif in combination with
996 the DISOPRED (62) algorithm as a comparison to PONDR, which produced very similar
997 results (online tool <http://www.yeastrc.org/pdr/>).

998

999

1000 **Oligomerization state of Mig1-GFP in the ‘pool’**

1001 Experimental *in vitro* assays of surface immobilized GFP coupled to simulations trained on
1002 these single-molecule intensity measurements but using noise levels comparable to *in vivo*
1003 cellular imaging conditions (27) indicate single-molecule sensitivity of GFP detection under
1004 our millisecond imaging conditions. However, if the nearest neighbor separation of individual
1005 GFP ‘foci’ are less than the optical resolution limit w of our microscope (which we measure
1006 as ~ 230 nm for GFP imaging) then distinct fluorescent foci will not be detected and instead
1007 will be manifest as a diffusive ‘pool’.

1008 If each GFP ‘foci’ in the pool has a mean stoichiometry S then the mean number of GFP
1009 foci, F , in the pool is n_{pool}/S and the ‘pool’ condition for nearest neighbor foci separation s
1010 indicates that $s < w$.

1011 The estimated range of mean total pool copy number from nucleus and cytoplasm
1012 combined, n_{pool} , is ~ 590 -1,100 molecules depending on extracellular glucose conditions.
1013 Approximating the cell volume as equal to the combined volumes of all uniformly separated
1014 foci in the pool (equal to the total number of foci multiplied by the volume of an equivalent
1015 sphere of radius r) indicates that $F \cdot 4\pi r^3/3 = 4\pi d^3/3$, thus $r = d/F^{1/3}$, where we use the mean
1016 measured cell diameter d of ~ 5 μm .

1017 However, mobile foci with a microscopic diffusion coefficient D will diffuse a mean
1018 two-dimensional distance b in focal plane of $(4D \cdot \Delta t)^{1/2}$ in a camera sampling time window Δt
1019 of 5 ms. Using $D \sim 6 \mu\text{m}^2 \text{s}^{-1}$ as a lower limit based on the measured diffusion of low
1020 stoichiometry cytoplasmic Mig1-GFP foci detected indicates $b \sim 340$ nm so the movement-
1021 corrected estimate for s is $r - b$, thus $s < w$ indicates that $r < b + w$, or $d/F^{1/3} < b + w$.

1022 Therefore, $d(S/n_{pool})^{1/3} < b + w$, and $S < n_{pool}((b+w)/d)^3$. Using ~ 590 -1,100 molecules from
1023 the measured mean range of n_{pool} indicates that the upper limit for S is in the range 0.8-1.4; in
1024 other words, Mig1-GFP foci in the pool are consistent with being a monomer.

1025

1026 **Analysis of the mobility of foci**

1027 For each accepted foci track the mean square displacement (MSD) was calculated from the
1028 optimized intensity centroid at time t of $(x(t), y(t))$ assuming a tracks of N consecutive image
1029 frames at a time interval $\tau = n\Delta t$ is (86, 87) where n is a positive integer is:

1030

$$1031 \quad \text{MSD}(\tau) = \text{MSD}(n\Delta t) = \frac{1}{N-1-n} \sum_{i=1}^{N-1-n} \left\{ [x(i\Delta t + n\Delta t) - x(i\Delta t)]^2 + [y(i\Delta t + n\Delta t) - y(i\Delta t)]^2 \right\}$$
$$= 4D\tau + 4\sigma^2$$

1032

1033 Here σ is the lateral (xy) localization precision which we estimate as ~ 40 nm (27). The
1034 microscopic diffusion coefficient D was then estimated from the gradient of a linear fit to the
1035 first four time interval data points of the MSD vs τ relation for each accepted foci track.

1036 To determine the proportion of mobile and immobile Mig1-GFP fluorescent foci we
1037 adapted an approach based on cumulative probability-distance distribution analysis (12).
1038 Here we generated cumulative distribution functions (CDFs) for all nuclear and cytoplasmic
1039 tracks, such that the CDF in each dataset is the probability distribution function p_c associated
1040 with r^2 , the square of the displacement between the first and second data points in each single
1041 track, which was generated for each track by calculating the proportion of all tracks in a
1042 dataset which have a value of r^2 less than that measured for that one track. The simplest CDF
1043 model assumes a Brownian diffusion propagator function $f(r^2)$ for a single effective diffusion
1044 coefficient component of:

1045

1046
$$f(r^2) = \frac{1}{4\pi D\Delta t} \exp\left(-\frac{r^2}{4D\Delta t}\right)$$

1047

1048 Here, D is the effective diffusion coefficient and Δt is image sampling time per frame (i.e. in
1049 our case 5 ms). This gives a CDF single component solution of the form:

1050

1051
$$p_c(r^2) = 1 - \exp\left(-\frac{r^2}{4D\Delta t}\right)$$

1052

1053 We investigated both single and more complex multi-component CDF models using either
1054 1,2 or 3 different D values in a weighted sum model of:

1055

1056
$$p_c(r^2) = \sum_{i=1}^n A_i \left(1 - \exp\left(-\frac{r^2}{4D_i\Delta t}\right)\right)$$

1057

1058 Here n is 1, 2 or 3. Multi-component fits were only chosen if they lowered the reduced χ^2 by
1059 $>10\%$. For cytoplasmic foci at *glucose* (+/-) and for nuclear foci at *glucose* (-) this indicated
1060 single component fits for diffusion coefficient with a D of $\sim 1-2 \mu\text{m}^2/\text{s}$, whereas nuclear foci
1061 at *glucose* (+) were fitted using two components of D , $\sim 20\%$ with a relatively immobile
1062 component, $D \sim 0.1-0.2 \mu\text{m}^2/\text{s}$, and the remainder a relatively mobile component, D
1063 $\sim 1-2 \mu\text{m}^2/\text{s}$, while using three components produced no statistically significant improvement
1064 to the fits. These values of D agreed to within experimental error to those obtained using a
1065 different method which fitted two analytical Gamma functions to the distribution of all
1066 calculated microscopic diffusion coefficients of tracked foci in the nucleus at *glucose* (+),
1067 which assumed a total probability distribution function p_γ of the form: (29)

1068

1069
$$p_\gamma(x, D) = \sum_{i=1}^2 \frac{A_i (m/D)^m x^{m-1} \exp(-mx/D)}{(m-1)!}$$

1070

1071 Here, m is the number of steps in the MSD vs τ trace for each foci track used to calculate D
1072 (i.e. in our instance $m=4$).

1073

1074 We also probed longer time scale effects on foci mobility for each accepted foci
1075 trajectory. Here, average MSD values were generated by calculating mean MSD values for
1076 corresponding time interval values across all foci trajectories in each dataset, but pooling
1077 traces into low stoichiometry (≤ 20 Mig1-GFP molecules per foci) and high stoichiometry ($>$
1078 20 Mig1-GFP molecules per foci). We compared different diffusion models over a 30 ms
1079 time interval scale, corresponding to the shortest time interval range from any of the mean
1080 MSD trace datasets.

1080

1081 We found in all cases that mean MSD traces could be fitted well (χ^2 values in the
1082 range 1-12) using a subdiffusion model of precision-corrected MSD = $4\sigma^2 + 4K\tau^\alpha$, where α
1083 the anomalous diffusion parameter and K is the transport parameter, analogous to the
1084 diffusion coefficient D in pure Brownian diffusion. Optimized fits indicated values of K in
1085 the range $0.08-0.2 \mu\text{m}^2/\text{s}$ and those for α of $\sim 0.4-0.8$. Corresponding fits to a purely Brownian
1086 diffusion model (i.e. $\alpha = 1$) generated much poorer fits (χ^2 values in the range 4-90).

1086

1087 We used both short timescale CDF analysis and longer timescale MSD analysis of Mig1
1088 tracks to try to gain as complete a picture of Mig1 mobility as possible. Short timescales

1088 avoid bias from photobleaching and diffusion out of the focal plane but longer timescales
1089 sample more of the cellular environment.
1090

1091 **Analyzing trans-nuclear tracks**

1092 The segmentation boundary output for the nucleus was fitted with a smoothing spline
1093 function, with smoothing parameter $p = 0.9992$ to sub-pixel precision. Trajectories which
1094 contained points on either side of the nuclear boundary were considered trans-nuclear. The
1095 crossing point on the nuclear boundary was found by linearly interpolating between the first
1096 pair of points either side of the nuclear boundary. Coordinates were normalized to this point
1097 and the crossing time and were rotated such that y' and x' lie perpendicular and parallel to the
1098 membrane crossing point.
1099

1100 **Investigating Mig1-GFP molecular turnover**

1101 Turnover of Mig1-GFP was investigated using fluorescence recovery after photobleaching
1102 (FRAP). In brief a 200 ms 10mW focused laser beam pulse of lateral width $\sim 1 \mu\text{m}$ was used
1103 to photobleach the fluorescently-labelled nuclear contents on a cell-by-cell basis and then \leq
1104 10 Slimfield images were recorded over different timescales spanning a range from 100 ms to
1105 $\sim 1,000$ s. The copy number of pool and foci in each image at subsequent time points t post
1106 focused laser bleach was determined as described and corrected for photobleaching. These
1107 post-bleach photoactive Mig1-GFP copy number values $C(t)$ could then be fitted using a
1108 single exponential recovery function:
1109

$$1110 \quad C(t) = C(0)(1 - \exp(-t/t_R))$$

1111

1112 Where t_R is the characteristic recovery (i.e. turnover) time (20). These indicated a value of
1113 133 ± 20 s (\pm SEM) for nuclear foci at glucose (+), and 3 ± 14 s for nuclear pool at
1114 *glucose* (+) and (-).
1115

1116 **Modeling the effective diameter of clusters**

1117 The effective diameter d of a cluster was estimated from the measured point spread function
1118 width $p_{f_{foci}}$ (defined at twice sigma value of the equivalent Gaussian fit from our single
1119 particle tracking algorithm) corrected for the blur due to particle diffusion in the camera
1120 exposure time of Δt as:
1121

$$1122 \quad d = p_{f_{foci}} - p_{GFP} - \sqrt{4D\Delta t}$$

1123

1124 Where D is the measured microscopic diffusion coefficient for that track and p_{GFP} is the
1125 measured point spread function width of surface-immobilized GFP (i.e. twice the sigma
1126 width of 230nm measured in our microscope, or 460nm). We explored a heuristic packing
1127 model of $d \sim S^a$ for Mig1-GFP monomers in each cluster, such that a tightly packed spherical
1128 cluster of volume V composed of S smaller ca. spherical monomers each of volume V_1 and
1129 diameter d_1 varied as $V = S.V_1$ thus $4\pi(d/2)^3 = S.4\pi(d_1/2)^3$, thus in the specific instance of a
1130 perfect spherical cluster model $a = 1/3$.

1131 In principle, for general shapes of clusters for different packing conformations we
1132 expect $0 \leq a \leq 1$ such that e.g. if clusters pack as a long, thin rod of Mig1 monomers which
1133 rotates isotropically during time Δt , then $a = 1$. Whereas, if Mig1 monomers bind to a
1134 putative additional 'anchor' type structure to occupy available binding sites in forming a
1135 cluster, such that the size of the cluster does not significantly change with S but is dependent

1136 on the size of the putative anchor structure itself, then $a = 0$. Our optimized fits indicate
1137 $a = 0.32 \pm 0.06$ (\pm SEM), i.e. consistent with an approximate spherical shape cluster model.

1138

1139 **Modeling the probability of overlap in *in vitro* fluorescent protein characterization**

1140 The probability that two or more fluorescent protein foci are within the diffraction limit of
1141 our microscope in the *in vitro* characterization assays was determined using a previously
1142 reported Poisson model (26) to be $\sim 10\%$ at the *in vitro* protein concentrations used here. Such
1143 overlapping fluorescent proteins are detected as higher apparent stoichiometry foci.

1144

1145 **PP7 RNA labelling and overlap integral**

1146 Similar Slimfield microfluidics experiments were performed on Mig1-mCherry and Mig1-
1147 mCherry Δ Znf strains containing 24 transcriptional reporter PP7 markers on the GYS1 gene
1148 and transformed with plasmids for the PP7 protein labelled with 2 GFPs. Mig1 foci are
1149 present at *glucose* (+) and upon switching to *glucose* (-) PP7 foci appear in similar locations
1150 to the Mig1 foci. Although Mig1 foci are mobile, the microscopic diffusion coefficient D for
1151 immobilized Mig1 is a putative overestimate for the equivalent D of the underlying target
1152 gene loci, $0.15 \mu\text{m}^2/\text{s}$ from CDF. The plateau of the MSD vs tau plot in Figure. 3B gives an
1153 estimate of the gene loci mobility range in space (although still an overestimate) and is
1154 $\sim 0.05 \mu\text{m}^2$. The square root of this is less than PSF width, and so colocalization between
1155 Mig1 and PP7 foci is expected.

1156

1157 The extent of colocalization between Mig1-mCherry and PP7-GFP detected foci was
1158 determined by calculating the overlap integral between each pair, whose centroids were
1159 within 5 pixels of each other. Assuming two normalized, 2D Gaussian intensity distributions
1160 g_1 and g_2 , for green and red foci respectively, centered around (x_1, y_1) with sigma width σ_1 ,
1161 and around (x_2, y_2) with width σ_2 , the overlap integral v is analytically determined as:

1162

$$v = \exp\left(-\frac{\Delta r^2}{2(\sigma_1^2 + \sigma_2^2)}\right)$$

1163

1164 Where

1165

$$1166 \Delta r^2 = (x_1^2 - x_2^2)^2 + (y_1^2 - y_2^2)^2$$

1167

1168 Previous studies have used an overlap integral of over 0.75 as a criteria for colocalization
1169 (26).

1170

1171 **Software and DNA sequence access**

1172 All our bespoke software developed, and Mig1 secondary structure prediction 3D coordinates
1173 `pymolMig1.pdb`, are freely and openly accessible via <https://sourceforge.net/projects/york-biophysics/>. The bespoke plasmid sequence information for the GFP reporter is openly
1174 accessible via <https://www.addgene.org/75360/>.

1175

1176 **Statistical tests and replicates**

1177 All statistical tests used are two-sided unless stated otherwise. For Slimfield imaging each
1178 cell can be defined as a biological replicate sampled from the cell population. We chose
1179 sample sizes of at least 30 cells which generated reasonable estimates for the sampled
1180 stoichiometry distributions, similar to those of previous *in vivo* Slimfield studies (20).

1182 Technical replicates are not possible with the irreversible photobleaching assay, however, the

1183 noise in all light microscopy experiments has been independently characterized for the
1184 imaging system used previously (27).
1185

1186 **Acknowledgments**

1187 Supported by the Biological Physical Sciences Institute, Royal Society, MRC (grant
1188 MR/K01580X/1), BBSRC (grant BB/N006453/1), Swedish Research Council and European
1189 Commission via Marie Curie-Network for Initial training ISOLATE (Grant agreement nr:
1190 289995) and the Marie Curie Alumni Association. We thank Magnus Alm Rosenblad and
1191 Sarah Shammass for assistance with RNABOB and PONDR, Marija Cvijovic and Michael
1192 Law for help with qPCR data analysis, Andrew Leech and Meg Stark for help with CD and
1193 EM. Thanks to Mark Johnston (CU Denver) for donation of Mig1 phosphorylation mutant
1194 plasmid, and Michael Elowitz (Caltech) for donation of the Mig1/Msn2/PP7 and Zn finger
1195 deletion strain.

1196

1197 **Conflict of interest**

1198 All the authors declare that they have no conflict of interests.

1199

1200

1201

1202 **References**

- 1203 1. F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins. *J.*
1204 *Mol. Biol.* **3**, 318–356 (1961).
- 1205 2. J. Gertz, E. D. Siggia, B. A. Cohen, Analysis of combinatorial cis-regulation in
1206 synthetic and genomic promoters. *Nature*. **457**, 215–8 (2009).
- 1207 3. O. G. Berg, R. B. Winter, P. H. Von Hippel, Diffusion-driven mechanisms of protein
1208 translocation on nucleic acids. 1. Models and theory. *Biochemistry*. **20**, 6929–6948
1209 (1981).
- 1210 4. A. Mahmutovic, O. G. Berg, J. Elf, What matters for lac repressor search in vivo--
1211 sliding, hopping, intersegment transfer, crowding on DNA or recognition? *Nucleic*
1212 *Acids Res.* **43**, 3454–64 (2015).
- 1213 5. S. E. Halford, J. F. Marko, How do site-specific DNA-binding proteins find their
1214 targets? *Nucleic Acids Res.* **32**, 3040–52 (2004).
- 1215 6. D. M. Gowers, S. E. Halford, Protein motion from non-specific to specific DNA by
1216 three-dimensional routes aided by supercoiling. *EMBO J.* **22**, 1410–8 (2003).
- 1217 7. S. T. Whiteside, S. Goodbourn, Signal transduction and nuclear targeting: regulation of
1218 transcription factor activity by subcellular localisation. *J. Cell Sci.* **104** (Pt 4, 949–55
1219 (1993).
- 1220 8. C. T. Harbison *et al.*, Transcriptional regulatory code of a eukaryotic genome. *Nature*.
1221 **431**, 99–104 (2004).
- 1222 9. H. G. Schmidt, S. Sewitz, S. S. Andrews, K. Lipkow, An Integrated Model of
1223 Transcription Factor Diffusion Shows the Importance of Intersegmental Transfer and
1224 Quaternary Protein Structure for Target Site Finding. **9** (2014),
1225 doi:10.1371/journal.pone.0108575.
- 1226 10. G.-W. Li, X. S. Xie, Central dogma at the single-molecule level in living cells. *Nature*.
1227 **475**, 308–15 (2011).
- 1228 11. P. Hammar *et al.*, The lac repressor displays facilitated diffusion in living cells.
1229 *Science*. **336**, 1595–8 (2012).
- 1230 12. J. C. M. Gebhardt *et al.*, Single-molecule imaging of transcription factor binding to
1231 DNA in live mammalian cells. *Nat. Methods*. **10**, 421–6 (2013).
- 1232 13. J. O. Nehlin, M. Carlberg, H. Ronne, Control of yeast GAL genes by MIG1 repressor:
1233 a transcriptional cascade in the glucose response. *EMBO J.* **10**, 3373–7 (1991).
- 1234 14. E. Frolova, Binding of the glucose-dependent Mig1p repressor to the GAL1 and GAL4
1235 promoters in vivo: regulation by glucose and chromatin structure. *Nucleic Acids Res.*
1236 **27**, 1350–1358 (1999).
- 1237 15. M. J. De Vit, J. a Waddle, M. Johnston, Regulated nuclear translocation of the Mig1
1238 glucose repressor. *Mol. Biol. Cell.* **8**, 1603–18 (1997).
- 1239 16. L. Bendrioua *et al.*, Yeast AMP-activated Protein Kinase Monitors Glucose
1240 Concentration Changes and Absolute Glucose Levels. *J. Biol. Chem.* **289**, 12863–75
1241 (2014).
- 1242 17. S. Shashkova, A. J. M. Wollman, M. C. Leake, S. Hohmann, The yeast Mig1
1243 transcriptional repressor is dephosphorylated by glucose-dependent and -independent
1244 mechanisms. *FEMS Microbiol. Lett.* **364** (2017), doi:10.1093/femsle/fnx133.

- 1245 18. Y. Lin, C. H. Sohn, C. K. Dalal, L. Cai, M. B. Elowitz, Combinatorial gene regulation
1246 by modulation of relative pulse timing. *Nature*. **527**, 54–58 (2015).
- 1247 19. M. Plank, G. H. Wadhams, M. C. Leake, Millisecond timescale slimfield imaging and
1248 automated quantification of single fluorescent protein molecules for use in probing
1249 complex biological processes. *Integr. Biol. (Camb)*. **1**, 602–12 (2009).
- 1250 20. R. Reyes-Lamothe, D. J. Sherratt, M. C. Leake, Stoichiometry and architecture of
1251 active DNA replication machinery in *Escherichia coli*. *Science*. **328**, 498–501 (2010).
- 1252 21. A. Badrinarayanan, R. Reyes-Lamothe, S. Uphoff, M. C. Leake, D. J. Sherratt, In vivo
1253 architecture and action of bacterial structural maintenance of chromosome proteins.
1254 *Science*. **338**, 528–31 (2012).
- 1255 22. M. C. Leake *et al.*, Stoichiometry and turnover in single, functioning membrane
1256 protein complexes. *Nature*. **443**, 355–8 (2006).
- 1257 23. D. A. Zacharias, J. D. Violin, A. C. Newton, R. Y. Tsien, Partitioning of lipid-
1258 modified monomeric GFPs into membrane microdomains of live cells. *Science*. **296**,
1259 913–6 (2002).
- 1260 24. S. A. McKinney, C. S. Murphy, K. L. Hazelwood, M. W. Davidson, L. L. Looger, A
1261 bright and photostable photoconvertible fluorescent protein. *Nat. Methods*. **6**, 131–3
1262 (2009).
- 1263 25. H. Miller, Z. Zhou, A. J. M. Wollman, M. C. Leake, Superresolution imaging of single
1264 DNA molecules using stochastic photoblinking of minor groove and intercalating
1265 dyes. *Methods*. **88**, 81–8 (2015).
- 1266 26. I. Llorente-Garcia *et al.*, Single-molecule in vivo imaging of bacterial respiratory
1267 complexes indicates delocalized oxidative phosphorylation. *Biochim. Biophys. Acta*.
1268 **1837**, 811–824 (2014).
- 1269 27. A. J. M. Wollman, M. C. Leake, Millisecond single-molecule localization microscopy
1270 combined with convolution analysis and automated image segmentation to determine
1271 protein concentrations in complexly structured, functional cells, one cell at a time.
1272 *Faraday Discuss.* **184**, 401–24 (2015).
- 1273 28. M. J. DeVit, M. Johnston, The nuclear exportin Msn5 is required for nuclear export of
1274 the Mig1 glucose repressor of *Saccharomyces cerevisiae*. *Curr. Biol.* **9**, 1231–1241
1275 (1999).
- 1276 29. M. Stracy *et al.*, Live-cell superresolution microscopy reveals the organization of RNA
1277 polymerase in the bacterial nucleoid. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E4390-9
1278 (2015).
- 1279 30. I. Izeddin *et al.*, Single-molecule tracking in live cells reveals distinct target-search
1280 strategies of transcription factors in the nucleus. *Elife*. **3** (2014),
1281 doi:10.7554/eLife.02230.
- 1282 31. S. A. Adam, The nuclear pore complex. *Genome Biol.* **2** (2001) REVIEWS007.
- 1283 32. C. Strambio-De-Castillia, M. Niepel, M. P. Rout, The nuclear pore complex: bridging
1284 nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 490–501 (2010).
- 1285 33. W. Yang, J. Gelles, S. M. Musser, Imaging of single-molecule translocation through
1286 nuclear pore complexes. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12887–12892 (2004).
- 1287 34. A. R. Lowe *et al.*, Selectivity mechanism of the nuclear pore complex characterized by

- 1288 single cargo tracking. *Nature*. **467**, 600–603 (2010).
- 1289 35. M. A. Treitel, M. Carlson, Repression by SSN6-TUP1 is directed by MIG1, a
1290 repressor/activator protein. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3132–6 (1995).
- 1291 36. Y. Phillip, G. Schreiber, Formation of protein complexes in crowded environments--
1292 from in vitro to in vivo. *FEBS Lett.* **587**, 1046–52 (2013).
- 1293 37. M. Lundin, J. O. Nehlin, H. Ronne, Importance of a flanking AT-rich region in target
1294 site recognition by the GC box-binding zinc finger protein MIG1. *Mol. Cell. Biol.* **14**,
1295 1979–1985 (1994).
- 1296 38. Z. Duan *et al.*, A three-dimensional model of the yeast genome. *Nature*. **465**, 363–7
1297 (2010).
- 1298 39. I. Unnikrishnan, S. Miller, M. Meinke, D. C. LaPorte, Multiple positive and negative
1299 elements involved in the regulation of expression of GSY1 in *Saccharomyces*
1300 *cerevisiae*. *J. Biol. Chem.* **278**, 26450–7 (2003).
- 1301 40. V. N. Uversky, I. M. Kuznetsova, K. K. Turoverov, B. Zaslavsky, Intrinsically
1302 disordered proteins as crucial constituents of cellular aqueous two phase systems and
1303 coacervates. *FEBS Lett.* **589**, 15–22 (2015).
- 1304 41. J. A. Toretsky, P. E. Wright, Assemblages: functional units formed by cellular phase
1305 separation. *J. Cell Biol.* **206**, 579–88 (2014).
- 1306 42. D. S. Goodsell, A. J. Olson, Structural Symmetry and Protein Function. *Annu. Rev.*
1307 *Biophys. Biomol. Struct.* **29**, 105–153 (2000).
- 1308 43. J. H. Fong *et al.*, Intrinsic Disorder in Protein Interactions: Insights From a
1309 Comprehensive Structural Analysis. *PLoS Comput. Biol.* **5**, e1000316 (2009).
- 1310 44. K. Sode, S. Ochiai, N. Kobayashi, E. Usuzaka, Effect of reparation of repeat sequences
1311 in the human alpha-synuclein on fibrillation ability. *Int. J. Biol. Sci.* **3**, 1–7 (2007).
- 1312 45. C. Avitabile *et al.*, Circular Dichroism studies on the interactions of antimicrobial
1313 peptides with bacterial cells. *Sci. Rep.* **4**, 337–360 (2014).
- 1314 46. H. Nishi, J. H. Fong, C. Chang, S. A. Teichmann, A. R. Panchenko, Regulation of
1315 protein-protein binding by coupling between phosphorylation and intrinsic disorder:
1316 analysis of human protein complexes. *Mol. Biosyst.* **9**, 1620–6 (2013).
- 1317 47. D. Mazza, A. Abernathy, N. Golob, T. Morisaki, J. G. McNally, A benchmark for
1318 chromatin binding measurements in live cells. *Nucleic Acids Res.* **40**, e119–e119
1319 (2012).
- 1320 48. D. Normanno *et al.*, Probing the target search of DNA-binding proteins in mammalian
1321 cells using TetR as model searcher. *Nat. Commun.* **6**, 7357 (2015).
- 1322 49. P. G. Needham, R. J. Trumbly, In vitro characterization of the Mig1 repressor from
1323 *Saccharomyces cerevisiae* reveals evidence for monomeric and higher molecular
1324 weight forms. *Yeast.* **23**, 1151–66 (2006).
- 1325 50. I. I. Cisse *et al.*, Real-Time Dynamics of RNA Polymerase II Clustering in Live
1326 Human Cells. *Science (80-.)*. **341**, 664–667 (2013).
- 1327 51. W.-K. Cho *et al.*, RNA Polymerase II cluster dynamics predict mRNA output in living
1328 cells. *Elife.* **5** (2016), doi:10.7554/eLife.13617.
- 1329 52. J. Qian *et al.*, B Cell Super-Enhancers and Regulatory Clusters Recruit AID

- 1330 Tumorigenic Activity. *Cell*. **159**, 1524–1537 (2014).
- 1331 53. M. Mir *et al.*, Dense Bicoid Hubs Accentuate Binding along the Morphogen Gradient.
1332 *bioRxiv* (2017).
- 1333 54. T. S. Karpova *et al.*, Concurrent fast and slow cycling of a transcriptional activator at
1334 an endogenous promoter. *Science*. **319**, 466–9 (2008).
- 1335 55. R. Rohs *et al.*, Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.*
1336 **79**, 233–69 (2010).
- 1337 56. J. Elf, G.-W. Li, X. S. Xie, Probing Transcription Factor Dynamics at the Single-
1338 Molecule Level in a Living Cell. *Science (80-.)*. **316**, 1191–1194 (2007).
- 1339 57. M. A. Lomholt, B. van den Broek, S.-M. J. Kalisch, G. J. L. Wuite, R. Metzler,
1340 Facilitated diffusion with DNA coiling. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8204–8
1341 (2009).
- 1342 58. B. van den Broek, M. A. Lomholt, S.-M. J. Kalisch, R. Metzler, G. J. L. Wuite, How
1343 DNA coiling enhances target localization by proteins. *Proc. Natl. Acad. Sci. U. S. A.*
1344 **105**, 15738–42 (2008).
- 1345 59. A. Esadze, J. Iwahara, Stopped-Flow Fluorescence Kinetic Study of Protein Sliding
1346 and Intersegment Transfer in the Target DNA Search Process. *J. Mol. Biol.* **426**, 230–
1347 244 (2014).
- 1348 60. D. M. Gowers, G. G. Wilson, S. E. Halford, Measurement of the contributions of 1D
1349 and 3D pathways to the translocation of a protein along DNA. *Proc. Natl. Acad. Sci.*
1350 *U. S. A.* **102**, 15883–8 (2005).
- 1351 61. B. Xue, A. K. Dunker, V. N. Uversky, Orderly order in protein intrinsic disorder
1352 distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J.*
1353 *Biomol. Struct. Dyn.* **30**, 137–49 (2012).
- 1354 62. J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, D. T. Jones, Prediction and
1355 functional analysis of native disorder in proteins from the three kingdoms of life. *J.*
1356 *Mol. Biol.* **337**, 635–45 (2004).
- 1357 63. G. Rivas, J. A. Fernández, A. P. Minton, Direct observation of the enhancement of
1358 noncooperative protein self-assembly by macromolecular crowding: indefinite linear
1359 self-association of bacterial cell division protein FtsZ. *Proc. Natl. Acad. Sci. U. S. A.*
1360 **98**, 3150–5 (2001).
- 1361 64. R. Babazadeh *et al.*, Osmostress-induced cell volume loss delays yeast Hog1 signaling
1362 by limiting diffusion processes and by Hog1-specific effects. *PLoS One*. **8**, e80901
1363 (2013).
- 1364 65. A. Miermont *et al.*, Severe osmotic compression triggers a slowdown of intracellular
1365 signaling, which can be explained by molecular crowding. *Proc. Natl. Acad. Sci. U. S.*
1366 *A.* **110**, 5725–30 (2013).
- 1367 66. J. Ma, W. Yang, Three-dimensional distribution of transient interactions in the nuclear
1368 pore complex obtained from single-molecule snapshots. *Proc. Natl. Acad. Sci. U. S. A.*
1369 **107**, 7305–7310 (2010).
- 1370 67. I. Axelsson, Characterization of proteins and other macromolecules by agarose gel
1371 chromatography. *J. Chromatogr. A.* **152**, 21–32 (1978).
- 1372 68. M. A. Hink *et al.*, Structural dynamics of green fluorescent protein alone and fused

- 1373 with a single chain Fv protein. *J. Biol. Chem.* **275**, 17556–60 (2000).
- 1374 69. L.-F. Liang, X. Da, T.-S. Chen, Y.-H. Pei, [Nucleoplasmic viscosity of living cells
1375 investigated by fluorescence correlation spectroscopy]. *Guang Pu Xue Yu Guang Pu*
1376 *Fen Xi.* **29**, 459–62 (2009).
- 1377 70. H.-J. Schüller, Transcriptional control of nonfermentative metabolism in the yeast
1378 *Saccharomyces cerevisiae*. *Curr. Genet.* **43**, 139–60 (2003).
- 1379 71. J. Liu *et al.*, Intrinsic disorder in transcription factors. *Biochemistry.* **45**, 6873–88
1380 (2006).
- 1381 72. V. N. Uversky, V. B. Patel, Intrinsically disordered proteins and their (disordered)
1382 proteomes in neurodegenerative disorders (2015), doi:10.3389/fnagi.2015.00018.
- 1383 73. E. M. Rubenstein *et al.*, Access denied: Snf1 activation loop phosphorylation is
1384 controlled by availability of the phosphorylated threonine 210 to the PP1 phosphatase.
1385 *J. Biol. Chem.* **283**, 222–30 (2008).
- 1386 74. R. D. Gietz, R. H. Schiestl, Frozen competent yeast cells that can be transformed with
1387 high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 1–4
1388 (2007).
- 1389 75. J. Warringer *et al.*, Trait variation in yeast is defined by population history. *PLoS*
1390 *Genet.* **7**, e1002111 (2011).
- 1391 76. C. Geijer *et al.*, Initiation of the transcriptional response to hyperosmotic shock
1392 correlates with the potential for volume recovery. *FEBS J.* **280**, 3854–67 (2013).
- 1393 77. S. Shashkova, M. C. Leake, Single-molecule fluorescence microscopy review:
1394 shedding new light on old problems. *Biosci. Rep.* **37** (2017).
- 1395 78. A.-K. Gustavsson *et al.*, Sustained glycolytic oscillations in individual isolated yeast
1396 cells. *FEBS J.* **279**, 2837–47 (2012).
- 1397 79. E. M. Puchner, J. M. Walter, R. Kasper, B. Huang, W. A. Lim, Counting molecules in
1398 single organelles with superresolution microscopy allows tracking of the endosome
1399 maturation trajectory, doi:10.1073/pnas.1309676110.
- 1400 80. L. H. Hartwell, J. Culotti, B. Reid, Genetic control of the cell-division cycle in yeast. I.
1401 Detection of mutants. *Proc. Natl. Acad. Sci. U. S. A.* **66**, 352–9 (1970).
- 1402 81. A. Khmelinskii *et al.*, Tandem fluorescent protein timers for in vivo analysis of protein
1403 dynamics. *Nat. Biotechnol.* **30**, 708–14 (2012).
- 1404 82. K. T. Tokuyasu, A technique for ultracryotomy of cell suspensions and tissues. *J. Cell*
1405 *Biol.* **57**, 551–65 (1973).
- 1406 83. J. Griffith, M. Mari, A. De Mazière, F. Reggiori, A Cryosectioning Procedure for the
1407 Ultrastructural Analysis and the Immunogold Labelling of Yeast *Saccharomyces*
1408 *cerevisiae*. *Traffic.* **9**, 1060–1072 (2008).
- 1409 84. N. J. Riccitelli, A. Lupták, Computational discovery of folded RNA domains in
1410 genomes and in vitro selected libraries. *Methods.* **52**, 133–40 (2010).
- 1411 85. Z. Obradovic, K. Peng, S. Vucetic, P. Radivojac, A. K. Dunker, Exploiting
1412 heterogeneous sequence properties improves prediction of protein disorder. *Proteins.*
1413 **61 Suppl 7**, 176–82 (2005).
- 1414 86. D. Gross, W. Webb, in *Spectroscopic Membrane Probes*, L. Loew, Ed. (CRC Press,

- 1415 ed. 2, 1988), pp. 19–45.
- 1416 87. X. Michalet, Mean square displacement analysis of single-particle trajectories with
1417 localization error: Brownian motion in an isotropic medium. *Phys. Rev. E*. **82**, 41914
1418 (2010).
- 1419 88. M. C. Leake, D. Wilson, M. Gautel, R. M. Simmons, The elasticity of single titin
1420 molecules using a two-bead optical tweezers assay. *Biophys. J.* **87**, 1112–35 (2004).
- 1421 89. M. C. Leake, D. Wilson, B. Bullard, R. M. Simmons, M. R. Bubb, The elasticity of
1422 single kettin molecules using a two-bead laser-tweezers assay. *FEBS Lett.* **535** (2003),
1423 doi:10.1016/S0014-5793(02)03857-7.
- 1424 90. P. Rice, I. Longden, A. Bleasby, EMBOSS: the European Molecular Biology Open
1425 Software Suite. *Trends Genet.* **16**, 276–7 (2000).
- 1426 91. M. C. Schmidt, R. R. McCartney, beta-subunits of Snf1 kinase are required for kinase
1427 function and substrate definition. *EMBO J.* **19**, 4936–43 (2000).
- 1428 92. G. Berben, J. Dumont, V. Gilliquet, P. A. Bolle, F. Hilger, The YDp plasmids: a
1429 uniform set of vectors bearing versatile gene disruption cassettes for *Saccharomyces*
1430 *cerevisiae*. *Yeast.* **7**, 475–7 (1991).
- 1431

1432 **Legends**

1433

1434 **Table 1. *S. cerevisiae* cell strains and plasmids.** List of all strains and plasmids used in this
1435 study.

1436

1437 **Table 2. Copy number data.** Mean average and SD of copy number in pool and foci in each
1438 compartment.

1439 **Table 3. Foci tracking data.** Mean average, SD and mean number detected per cell (N) of
1440 stoichiometry values (molecules), and microscopic diffusion coefficients D in each
1441 compartment detected within the depth of field.

1442 **Table 4. *snf1Δ* foci tracking and copy number data.** Upper panel: Mean average, SD and
1443 mean number detected per cell (N) of stoichiometry values (molecules), and microscopic
1444 diffusion coefficients D in each compartment detected within the depth of field. Lower panel:
1445 Mean average and SD of copy number in pool and foci in each compartment.

1446 **Table 5. Number of potential Mig1 target promoter sites per chromosome.** List of
1447 *S.cerevisiae* chromosomes indicating the length of a chromosome, total number of potential
1448 Mig1 target sites identified and then the number of sites on promoters assuming a promoter
1449 region up to 500bp upstream of a gene.

1450 **Table 6. Bioinformatics analysis for intrinsically disordered sequences.** Predictions for
1451 the presence of intrinsically disordered sequences in Mig1, Msn2 and LacI, and of the
1452 positions of phosphorylation sites in Mig1 and Msn2.

1453 **Video 1. Dual-color fluorescence microscopy assay at *glucose* (+).** Example cell showing
1454 *glucose* (+) nuclear Mig1-GFP localization (green, distinct foci black arrows), Nrd1-mCherry
1455 (red) and segmented cell body (orange) and nuclear envelope (cyan) indicated, slowed 15x.

1456 **Video 2. Dual-color fluorescence microscopy assay at *glucose* (-).** Example cell showing
1457 *glucose* (-) Mig1-GFP localization (green, distinct foci black arrows), Nrd1-mCherry (red)
1458 and segmented cell body (orange) and nuclear envelope (cyan) indicated, slowed 200x.

1459

1460 **Figure. 1. Single-molecule Slimfield microscopy of live cells reveals Mig1 clusters. (A)**
1461 Dual-color fluorescence microscopy assay. Mig1-GFP localization change (cyan, right
1462 panels) depending on glucose availability. **(B)** Example Slimfield micrographs of change of
1463 Mig1-GFP localization (green) with glucose for three cells, nuclear Nrd1-mCherry indicated
1464 (red, left), mean and SEM errorbounds of total cytoplasmic (yellow) and nuclear (blue)
1465 contributions shown (lower panel), n=15 cells. Display scale fixed throughout each time
1466 course to show pool and foci fluorescence. **(C)** Example Slimfield micrographs of cells
1467 showing nuclear (left), trans-nuclear (center) and cytoplasmic (right) Mig1-GFP localization
1468 (green, distinct foci white arrows), Nrd1-mCherry (red) and segmented cell body (yellow)
1469 and nuclear envelope (blue) indicated. Display scales adjusted to only show foci. **(D)** Kernel
1470 density estimations (KDE) for Mig1-GFP content in pool and foci for cytoplasm and nucleus
1471 at *glucose* (+/-), n=30 cells.

1472

1473 **Figure. 1 – Figure Supplement 1. Brightfield and fluorescence micrographs of key**
1474 **strains and glucose conditions.**

1475 Representative Slimfield fluorescence images obtained from the strains and different
1476 extracellular glucose conditions used in this study. Brightfield non-fluorescence images,
1477 segmentation perimeter indicated for cell body (orange) and nucleus (cyan), and fluorescence
1478 images are indicated, the latter showing both green and red channels obtained as the frame
1479 average from the first five consecutive Slimfield images. Fluorescence images are of the first
1480 illuminated frame and are all normalized by total pixel intensity. Snf1AS represents an ATP
1481 analog sensitive version of Snf1, Snf1-I132G. For the Mig1-mEos2 strain (inset, bottom
1482 right) this shows the brightfield image (left panel), a 300 consecutive frame average from the
1483 red channel after photoconversion (middle panel) and super-resolution false color heat map
1484 reconstruction, 40nm lateral resolution, >2,000 localizations (right panel).

1485

1486 **Figure. 1 - Figure Supplement 2. Fluorescent reporter strains have similar viability to**
1487 **wild type, with relatively fast maturation of fluorescent protein, and no evidence for**
1488 **GFP-mediated oligomerization. (A)** (left panel) Mean doubling time \pm SEM, number of
1489 cultures n=6; (right panel) relative expression of *MIG1* to constitutive *ACT1* using qPCR in
1490 the wild type and Mig1-mGFP in cells pre-grown in elevated (4%) and depleted
1491 (0.2%) glucose, SD error bars, n=3 repeats for each. **(B)** ‘Monomeric’ mGFP (red) vs
1492 Standard enhanced eGFP (blue) *in vitro* intensity KDE distributions. GFP/mCherry
1493 maturation. n=1000 foci **(C)** After continuous illumination images were taken at subsequent
1494 time intervals. To prevent appearance of newly synthesized fluorescent proteins, 100 μ g/ml
1495 cycloheximide was added 1h prior to photobleaching. Upper panels represent
1496 autofluorescence appearance in green and red channels in BY4741 wild type cells. Standard
1497 epifluorescence images (green/red) overlaid on brightfield (gray). **(D)** GFP and mCherry
1498 maturation in minimal YNB media with complete amino acid supplement and 4% glucose.
1499 The background-corrected total cellular fluorescence intensity for the wild type
1500 (autofluorescence) and Mig1-GFP:Nrd1-mCherry strain was quantified at each time point for
1501 each cell in ImageJ. Error bars indicate SEM. **(E) and (F)** *In vivo* Mig1-GFP vs Mig1-mGFP
1502 stoichiometry KDE distributions compared in *glucose* (+) and *glucose* (-) respectively
1503 ($\chi^2=0.28$, $p=1 \times 10^{-132}$ and $\chi^2=0.011$, $p=3 \times 10^{-216}$). n=30 cells.

1504 **Figure. 1 - Figure Supplement 3. *In vivo* Mig1-GFP foci intensity traces as a function of**
1505 **time.** Showing the raw tracked intensity (blue line) and filtered (red squares) using an edge
1506 preserving Chung-Kennedy filter (88, 89) (15ms window) for (A) nuclear foci (B)
1507 cytoplasmic foci and (C) overtracked foci from >0.5s into the photobleach with steps at
1508 single GFP intensity. (D) The intensity distribution of Mig1-GFP *in vivo* at *glucose* (+/-)
1509 using kernel density estimate. Mode values are identical within error: 5200±1500, 4600±700
1510 respectively.

1511 **Figure. 2. Mig1 foci stoichiometry, mobility and localization depend on glucose.** Heat
1512 map showing dependence of stoichiometry of detected GFP-labeled Mig1 foci with *D* under
1513 (A) *glucose* (+) and (B) *glucose* (-) extracellular conditions. Mean values for *glucose* (+)
1514 nuclear and *glucose* (-) cytoplasmic foci indicated (arrows). n=30 cells. Heat maps generated
1515 using 1,000 square pixel grid and 15 pixel width Gaussians at each foci, using variable color
1516 scales specified by colorbar on the right.

1517 **Figure. 2 – Figure Supplement 1. Mig1 phosphorylation does not affect clustering but**
1518 **regulates localization.** Heat maps showing dependence of stoichiometry of detected GFP-
1519 labeled Mig1 foci with *D* in (A, B) *SNF1* deletion strain, (C) strain with ATP analog sensitive
1520 variant of Snf1, Snf1-I132G, in presence of 1NM-PP1, and (D, E) strain with four serine
1521 phosphorylation sites of Mig1 mutated to alanine.

1522
1523 **Figure. 2 – Figure Supplement 2. Wild type Snf1 and analog sensitive have similar effect**
1524 **on Mig1.** (A) Representative Slimfield fluorescence images showing expected Mig1
1525 localization for a strain carrying an analog sensitive Snf1 and upon treatment with DMSO.
1526 (B, C) Heat maps showing dependence of stoichiometry of detected GFP-labeled Mig1 foci
1527 with microscopic diffusion coefficient *D* in analog sensitive Snf1.

1528

1529 **Figure. 2 – Figure Supplement 3. Boxplot summary of wild type and mutant Mig1**
1530 **stoichiometry and microscopic diffusion coefficient.** Boxplots for stoichiometry (top) and
1531 diffusion coefficient (bottom). Median in red, quartiles in blue. Black dotted lines mark the
1532 extrema, limited to 4x and 2x upper quartile value for stoichiometry and microscopic
1533 diffusion coefficient *D* respectively with outliers marked with red crosses.

1534

1535 **Figure. 3. Repressor clusters have heterogeneous mobility depending on localization.**
1536 (A) Cumulative probability, *glucose* (+) nuclear tracks (blue) and 2 component exponential
1537 fit (red), with dual Gamma fit to *D* (inset) with similar parameters. (B) Mean MSD vs τ (i.e.
1538 time interval τ) from cytoplasmic (yellow), small (blue, stoichiometry ≤ 20 Mig1-GFP
1539 molecules) and large nuclear (purple, stoichiometry > 20 Mig1-GFP molecules) foci, SEM
1540 indicated, on log-log axes, n=30 cells for *glucose* (+) and (-). Anomalous diffusion model fits
1541 to time intervals ≤ 30 ms (dashed black line), anomalous coefficient $\alpha=0.4-0.8$. (C) Heat map
1542 of trans-nuclear track localizations normalized to crossover point, generated using 1000
1543 square pixel grid and 10 pixel width Gaussians at each localization (D) distance parallel (left)
1544 and perpendicular (right) to nuclear envelope with time, normalized to crossover point for
1545 Mig1-GFP foci entering (blue) and leaving the nucleus (red), (E) dwell times at nuclear

1546 envelope and single exponential fits (dotted). (F) Example *glucose* (+) single cell FRAP
1547 Slimfield images, fixed display scale (G) mean and SEM nuclear intensity after bleaching,
1548 $n=5$ and 7 cells for *glucose* (-/+), respectively.

1549 **Figure. 3 – Figure Supplement 1. Cumulative probability distance analysis reveals a**
1550 **single mobile population in the cytoplasm at glucose (+/-) and in the nucleus and glucose**
1551 **(-).** Cumulative density functions of first displacement in trajectories (blue) with appropriate
1552 fits (red). Bottom right panel indicates Mig1 mutant for which the Zn finger domain has been
1553 deleted.

1554 **Figure. 4. Mig1 clusters are stabilized by depletion forces and bind to promoter targets.**
1555 (A) Zoom-in on pairwise difference distribution for stoichiometry of Mig1-GFP foci, 7-mer
1556 intervals (dashed) and power spectrum (inset), mean and Gaussian sigma error (arrow). (B)
1557 Stoichiometry for Mig1-GFP clusters *in vitro* in PEG absence (blue)/presence (red). $n=1000$
1558 foci. Inset shows the full range while outer zooms in on cluster stoichiometry. (C) 3C model
1559 of chromosomal DNA (blue shaded differently for each chromosome) with overlaid Mig1
1560 promoter binding sites from bioinformatics (red), simulated image based on model with
1561 realistic signal and noise added (inset). (D) Cluster (red) and monomer (dark blue) model
1562 (goodness-of-fit $R^2 < 0$) for Mig1-GFP stoichiometry (10 replicates) compared against
1563 experimental data (cyan, $R^2 = 0.75$).
1564

1565 **Figure. 4 – Figure Supplement 1. Additional Mig1 cluster investigations.** (A) Zoom-in on
1566 pairwise difference distribution for stoichiometry of GFP-labeled Mig1 foci detected during
1567 FRAP, ~8-mer intervals (dashed lines) and power spectrum (inset) shown, mean and
1568 Gaussian sigma error (arrow). (B) GFP-labeled Mig1 cluster size as a function of
1569 stoichiometry with power law fit indicated. (C) Immuno-gold transmission electron
1570 microscopy for negatively stained 90nm cryosections of (upper panel) two different Mig1-
1571 GFP cells and (lower panel) two different Msn2-GFP cells, with zoom in (inset).
1572

1573 **Figure. 4 – Figure Supplement 2. *In vitro* cluster characterization.** (A) Native PAGE of
1574 total cell protein extracts obtained from cells grown in 4% *glucose* (+) and 0.05% *glucose* (-)
1575 conditions followed by western blotting and probing with anti-GFP antibodies. (B)
1576 Coomassie staining of purified Mig1-mGFP fraction indicates a single band that corresponds
1577 to the size of a Mig1-GFP monomer (molecular weight 83.4kDa). (C) Mig1 phosphorylation
1578 status is detected by SDS-PAGE on total cell protein extracts obtained from cells grown in
1579 different glucose conditions followed by western blotting. De/phosphorylation of Mig1 in
1580 *glucose* (+/-) respectively is not affected by the *SNF1- I132G* mutation. (D) Slimfield images
1581 of Mig1-GFP clusters *in vitro* in the absence (left) and presence of PEG (right), display
1582 intensity scaled in units of GFP per 5 pixel circular region of interest (ROI). (E) Distribution
1583 of stoichiometry for mGFP clusters *in vitro* in absence (blue)/presence (red) of PEG and the
1584 expected distribution of overlapping mGFP monomers (yellow).
1585

1586 **Figure. 4 – Figure Supplement 3. Additional 3C modelling.** (A) 3C model (blue) with
1587 overlaid bound Mig1 clusters to promoter binding sites from bioinformatics (red), and Mig1
1588 clusters near the NE (green); (B) predicted stoichiometry distributions for GFP-labeled Mig1

1589 foci in the nucleus at elevated extracellular glucose for a range of different binding models,
1590 including: a model which simulates both nuclear envelope (NE) translocating clusters and
1591 cluster binding to promoter targets (yellow), a model which simulates both nuclear envelope
1592 (NE) translocating monomers and monomer binding to promoter targets and DNA (blue), and
1593 a model which simulates just cluster binding to promoter targets but excludes any effects
1594 from translocating clusters (purple). These models are optimized to the experimentally
1595 determined stoichiometry distribution (cyan); (C) predicted Mig1 monomer stoichiometry
1596 distributions for Mig1 bound to promoter sites in three different orientations $\sim 10^\circ$ apart.

1597

1598 **Figure. 5. Msn2 and Mig1 forms functional clusters colocalized to transcribed mRNA**
1599 **from their target genes.** (A) Kernel density estimations for Msn2-GFP in pool and foci for
1600 cytoplasm and nucleus at *glucose* (+/-). (B) Heat maps showing dependence of stoichiometry
1601 and *D* of detected Msn2-GFP foci, n=30 cells. (C) Slimfield imaging on the same cell in
1602 which microfluidics is used to switch from *glucose* (+) to *glucose* (-) indicating the
1603 emergence of PP7-GFP foci at *glucose* (-) which are coincident with Mig1-mCherry foci at
1604 *glucose* (+), dependent on the Mig1 Zn finger (same intensity display scales throughout).
1605 These Mig1 and PP7 foci have a high level of colocalization as seen from (D) the distribution
1606 of the numerical overlap integral between foci in red and green channels at *glucose* (+) and
1607 *glucose* (-) respectively, peaking at ~ 0.95 . n=21 cells. (E) Two example cells showing at
1608 *glucose* (-) Msn2-mKO2 foci colocalize with PP7-GFP foci. PP7-2xGFP and Msn2-mKO2
1609 images are frame averages of $\sim 1,000$ frames, Mig1-mCherry is a Slimfield image.

1610

1611 **Figure. 6. Mig1 and Msn2 contain disordered sequences which may mediate cluster**
1612 **formation.** (A) Structural prediction for Mig1; Zn finger motif (cyan), disordered sections
1613 (red) from PyMOL, beta sheet (gray), phosphorylation sites (yellow); zoom-in indicates
1614 structure of conserved Zn finger from PSI-BLAST to PDB ID: 4R2E (Wilms tumor protein,
1615 WT1). (B) DISOPRED prediction for Mig1 and Msn2; disordered regions (red), Zn finger
1616 regions (cyan). (C) Circular dichroism of Mig1-GFP *in vitro* in PEG absence (blue)/presence
1617 (orange) (D) Distribution of nearest neighbor distances for Mig1 sites within promoters on
1618 same (blue) or different (red) chromosome. (E) Schematic of depletion-stabilized Mig1
1619 cluster bound to multiple promoter targets (Zn finger PDB ID: 4R2E). (F) Amino acid
1620 residue electrostatic charge plots for Mig1 and Msn2 from EMBOSS (90) Residues 'D' and 'E'
1621 are assigned a charge of -1, 'K' and 'R' a charge of +1, and the residue 'H' is assigned a charge
1622 of +0.5, then a rolling 75 amino acid residue window is used.

Figures and Tables

Strain name	Background	Genotype	Source/Reference
YSH1351	S288C	<i>MATa HIS3D0 LEU2D1 MET15D0 URA3D0</i>	S. Hohmann collection
YSH1703	W303-1A	<i>MATa mig1Δ::LEU2 snf1Δ::KanMX</i>	S. Hohmann collection
YSH2267	BY4741	<i>MATa his3D1 leu2D0 met15D0 ura3D0 mig1Δ::KanMX NRD1-mCherry-hphNT1</i>	S. Hohmann collection
YSH2350	BY4741	<i>MATa MSN2-GFP-HIS3 NRD1-mCherry-hphNT1 MET LYS</i>	(64)
YSH2856	BY4741	<i>MATa MIG1-eGFP-KanMX NRD1-mCherry-HphNT1 snf1Δ::LEU2 MET LYS</i>	This study
YSH2348	BY4741	<i>MATa MIG1-GFP-HIS3 NRD1-mCherry-hphNT1 MET LYS</i>	(16)
YSH2862	BY4741	<i>MATa MIG1-GFPmut3-HIS3</i>	This study
YSH2863	BY4741	<i>MATa MIG1-GFPmut3-HIS3 NRD1-mCherry-HphMX4</i>	This study
YSH2896	BY4741	<i>MATa MIG1-mEOs2-HIS3</i>	This study
ME404	BY4741	"BY4741 <i>MSN2-mKO2::LEU2 MIG1-mCherry::spHIS5 GSY1-24xPP7::KANMX msn4Δ mig2Δ nrg1::HPHMX nrg2::Met15 SUC2::NatMX</i> "	(18)
ME412	BY4741	BY4741 <i>MSN2-mKO2::LEU2 MIG1(Δaa36-91)-mCherry::spHIS5 GSY1-24xPP7::KANMX msn4Δ mig2Δnrg1::HPHMX nrg2::Met15</i>	(18)
ME411	BY4741	<i>MIG1(Δaa36-91)-mCherry::spHIS5 GSY1-24xPP7::KANMX msn4Δ mig2Δnrg1::HPHMX nrg2::Met15</i>	(18)

Plasmid name	Description	Source/Reference
pMIG1-HA	<i>HIS3</i>	(91)
pSNF1-TAP	<i>URA3, in pRS316</i>	S. Hohmann collection
pSNF1-I132G-TAP	<i>URA3, in pRS316</i>	S. Hohmann collection
pmGFPS	<i>HIS3, GFPmut3 S65G, S72A, A206K</i>	This study
pMig1-mGFP	<i>6xHIS-Mig1-GFPmut3 in pRSET A</i>	This study
pmEOs2	<i>mEOs2-HIS3 in pMK-RQ</i>	This study
YDp-L	<i>LEU2</i>	(92)
YDp-H	<i>HIS3</i>	(92)
BM3726	<i>Mig1 (Ser222,278,311,381 → Ala), URA3, in pRS316</i>	M. Johnston collection (28)
pDZ276	<i>PP7-2xGFP::URA3</i>	(18)

Table 1. *S. cerevisiae* cell strains and plasmids.

	Mig1-GFP				Msn2-GFP			
	Glucose (+)		Glucose (-)		Glucose (+)		Glucose (-)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cytoplasmic Pool	509	274	949	394	1422	977	2487	1360
Nuclear Pool	77	101	140	97	551	608	1692	1221
Total Pool	586	336	1088	392	1973	1585	4179	2581
Cytoplasmic Spots	57	79	311	212	333	196	776	635
Nuclear Spots	190	99	35	63	81	138	320	269
Total Spots	246	100	345	203	414	334	1096	904
Total Cytoplasm	580	276	1156	399	1755	1173	3263	1995
Total Nuclear	226	155	176	124	632	746	2012	1490
Total Cell	806	353	1331	352	2387	1919	5274	3485

Table 2. Copy number data.

	Mig1-GFP						Msn2-GFP					
	Glucose (+)			Glucose (-)			Glucose (+)			Glucose (-)		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
Stoichiometry of Nuclear Spots	19.0	16.2	7.2	8.5	4.8	5.8	34.5	26.6	3.5	46.5	31.6	4.7
Diffusion Constant of Nuclear Spots ($\mu\text{m}^2/\text{s}$)	0.8	0.8	7.2	1.3	1.5	5.8	0.7	0.9	3.5	0.9	0.9	4.7
Stoichiometry of Trans-Nuclear Spots	10.6	10.2	1.0	8.7	5.3	5.1	21.8	16.7	1.9	43.9	35.0	0.9
Diffusion Constant of Trans-Nuclear Spots ($\mu\text{m}^2/\text{s}$)	1.3	1.2	1.0	1.5	1.6	5.1	1.5	1.2	1.9	1.1	1.1	0.9
Stoichiometry of Cytoplasmic Spots	6.6	4.9	1.1	7.2	3.7	17.8	25.7	19.5	4.8	30.1	17.5	4.0
Diffusion Constant of Cytoplasmic Spots ($\mu\text{m}^2/\text{s}$)	1.4	1.4	1.1	1.2	1.2	17.8	1.2	1.1	4.8	1.0	1.4	4.0

Table 3. Foci tracking data.

	Mig1-GFP <i>snf1Δ</i>					
	Glucose (+)			Glucose (-)		
	Mean	SD	N	Mean	SD	N
Stoichiometry of Nuclear Spots	17.5	10.9	13.2	23.5	15.4	10.9
Diffusion Constant of Nuclear Spots ($\mu\text{m}^2/\text{s}$)	1.1	1.1	13.2	0.7	0.8	10.9
Stoichiometry of Trans-Nuclear Spots	8.9	6.0	1.2	12.7	6.1	0.5
Diffusion Constant of Trans-Nuclear Spots ($\mu\text{m}^2/\text{s}$)	1.9	2.0	1.2	1.1	1.4	0.5
Stoichiometry of Cytoplasmic Spots	6.2	2.2	5.0	8.3	4.1	9.1
Diffusion Constant of Cytoplasmic Spots ($\mu\text{m}^2/\text{s}$)	1.3	1.2	5.0	1.0	1.2	9.1
Copy Numbers						
Cytoplasmic Pool	947	728	30	608	450	30
Nuclear Pool	807	398	30	611	325	30
Total Pool	1754	1127	30	1219	775	30
Cytoplasmic Spots	118	169	30	334	374	30
Nuclear Spots	162	69	30	164	71	30
Total Spots	280	238	30	498	445	30
Total Cytoplasm	1065	897	30	941	824	30
Total Nuclear	969	467	30	775	396	30
Total Cell	2034	1364	30	1717	1220	30

Table 4. *snf1Δ* foci tracking and copy number data.

Chromosome	length (bp)	N sites identified	N promoter sites
I	230218	41	1
II	813184	134	10
III	316620	52	2
IV	1531933	240	14
V	576874	109	8
VI	270161	58	4
VII	1090940	168	13
VIII	562643	92	2
IX	439888	94	8
X	745751	125	6
XI	666816	117	6
XII	1078177	194	12
XIII	924431	157	6
XIV	784333	135	3
XV	1091291	185	11
XVI	948066	163	6

Table 5. Number of potential Mig1 target promoter sites per chromosome.

Msn2:

Predicted residues: 704	Number Disordered Regions: 12
Number residues disordered: 394	Longest Disordered Region: 145
Overall percent disordered: 55.97	Average Prediction Score: 0.5577
Predicted disorder segment [1]-[2]	Average Strength= 0.8759
Predicted disorder segment [16]-[33]	Average Strength= 0.6958
Predicted disorder segment [55]-[199]	Average Strength= 0.8311
Predicted disorder segment [222]-[249]	Average Strength= 0.8237
Predicted disorder segment [322]-[365]	Average Strength= 0.8820
Predicted disorder segment [410]-[428]	Average Strength= 0.7475
Predicted disorder segment [469]-[480]	Average Strength= 0.6545
Predicted disorder segment [510]-[549]	Average Strength= 0.8040
Predicted disorder segment [572]-[641]	Average Strength= 0.9319
Predicted disorder segment [660]-[667]	Average Strength= 0.6829
Predicted disorder segment [694]-[695]	Average Strength= 0.5325
Predicted disorder segment [699]-[704]	Average Strength= 0.6783

Mig1:

Predicted residues: 504	Number Disordered Regions: 9
Number residues disordered: 372	Longest Disordered Region: 95
Overall percent disordered: 73.81	Average Prediction Score: 0.7008
Predicted disorder segment [1]-[12]	Average Strength= 0.8252
Predicted disorder segment [25]-[33]	Average Strength= 0.6502
Predicted disorder segment [77]-[171]	Average Strength= 0.8758
Predicted disorder segment [173]-[240]	Average Strength= 0.9051
Predicted disorder segment [242]-[249]	Average Strength= 0.5554
Predicted disorder segment [254]-[272]	Average Strength= 0.7890
Predicted disorder segment [292]-[310]	Average Strength= 0.8225
Predicted disorder segment [327]-[386]	Average Strength= 0.8355
Predicted disorder segment [423]-[504]	Average Strength= 0.9136

LacI:

Predicted residues: 360	Number Disordered Regions: 8
Number residues disordered: 149	Longest Disordered Region: 48
Overall percent disordered: 41.39	Average Prediction Score: 0.4418
Predicted disorder segment [1]-[4]	Average Strength= 0.6245
Predicted disorder segment [18]-[52]	Average Strength= 0.6710
Predicted disorder segment [55]-[81]	Average Strength= 0.7443
Predicted disorder segment [88]-[100]	Average Strength= 0.5841
Predicted disorder segment [186]-[187]	Average Strength= 0.5429
Predicted disorder segment [238]-[256]	Average Strength= 0.6208
Predicted disorder segment [258]-[258]	Average Strength= 0.5028
Predicted disorder segment [313]-[360]	Average Strength= 0.8331

Phosphorylation sites of Mig1 and Msn2 (uniprot.org, accessed February, 2016):

Mig1 Phosphorylation site	Disorder segment	Msn2 Phosphorylation site	Disorder segment
S264	[254]-[272]	S194	[55]-[199]
S278	-	S201	-
T280	-	S288	-
S302	[292]-[310]	S304	-
S311	[292]-[310]	S306	-
S314	-	S308	-
S80	[77]-[171]	S432	-
S108	[77]-[171]	S451	-
S214	[173]-[240]	S582	[572]-[641]
S218	[173]-[240]	S620	[572]-[641]
S222	[173]-[240]	S625	[572]-[641]
S303	[292]-[310]	T627	[572]-[641]
S310	[292]-[310]	S629	[572]-[641]
S350	[327]-[386]	S633	[572]-[641]
S367	[327]-[386]		
S370	[327]-[386]		
T371	[327]-[386]		
S377	[327]-[386]		
S379	[327]-[386]		
S381	[327]-[386]		
S400	-		
S402	-		
T455	[423]-[504]		

Table 6. Bioinformatics analysis for intrinsically disordered sequences.

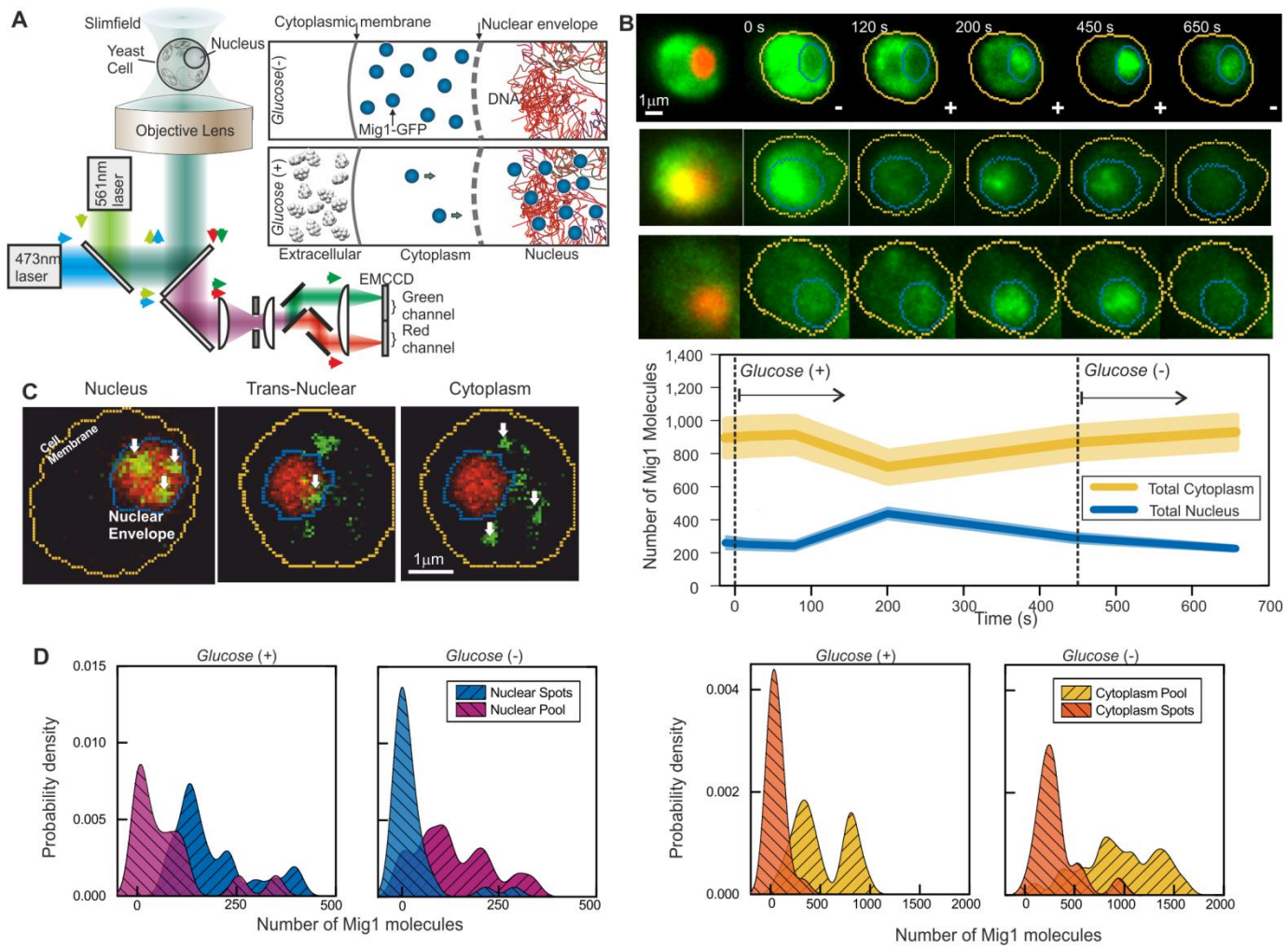


Figure 1. Single-molecule Slimfield microscopy of live cells reveals Mig1 clusters.

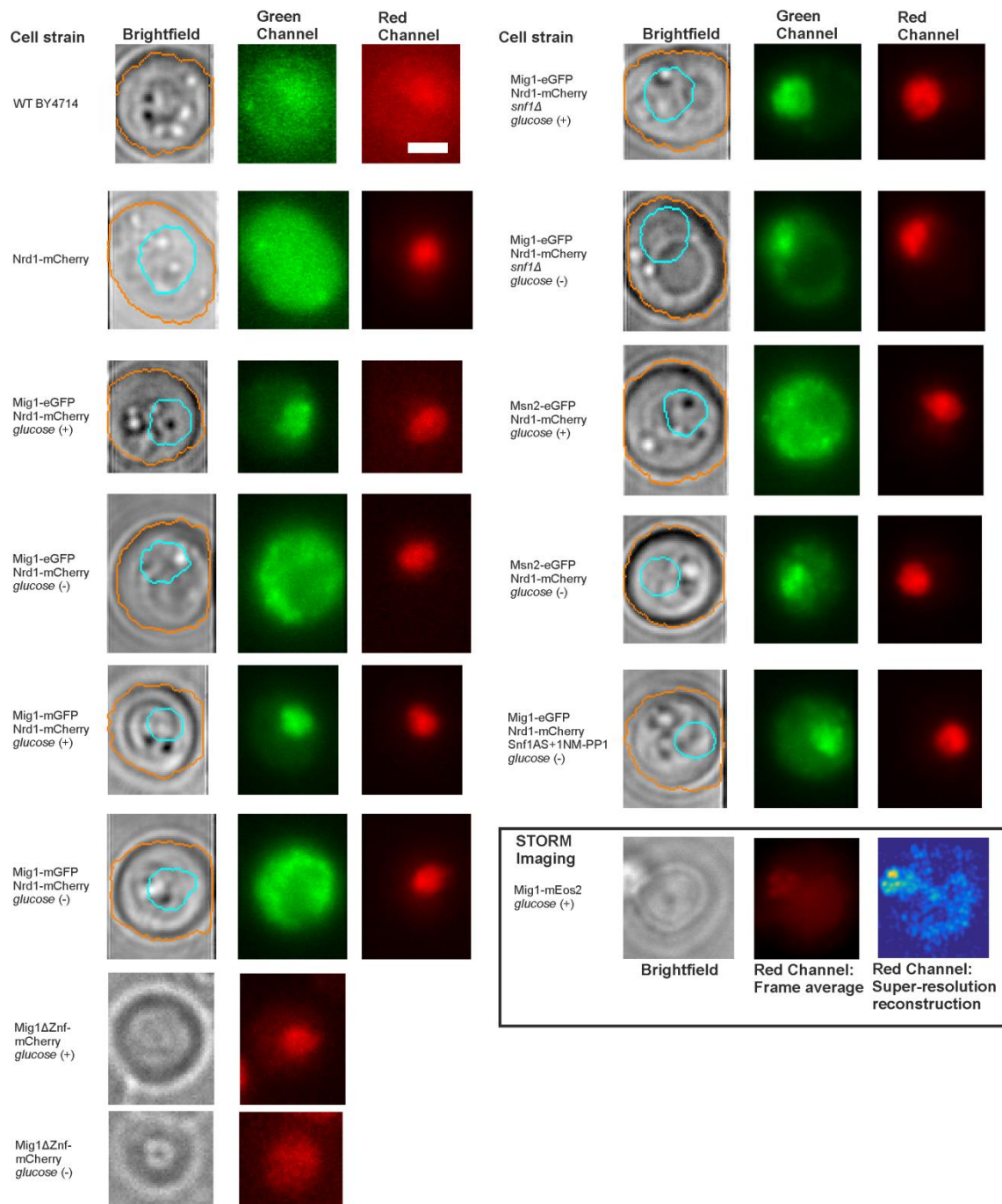


Figure. 1 – Figure Supplement 1. Brightfield and fluorescence micrographs of key strains and glucose conditions.

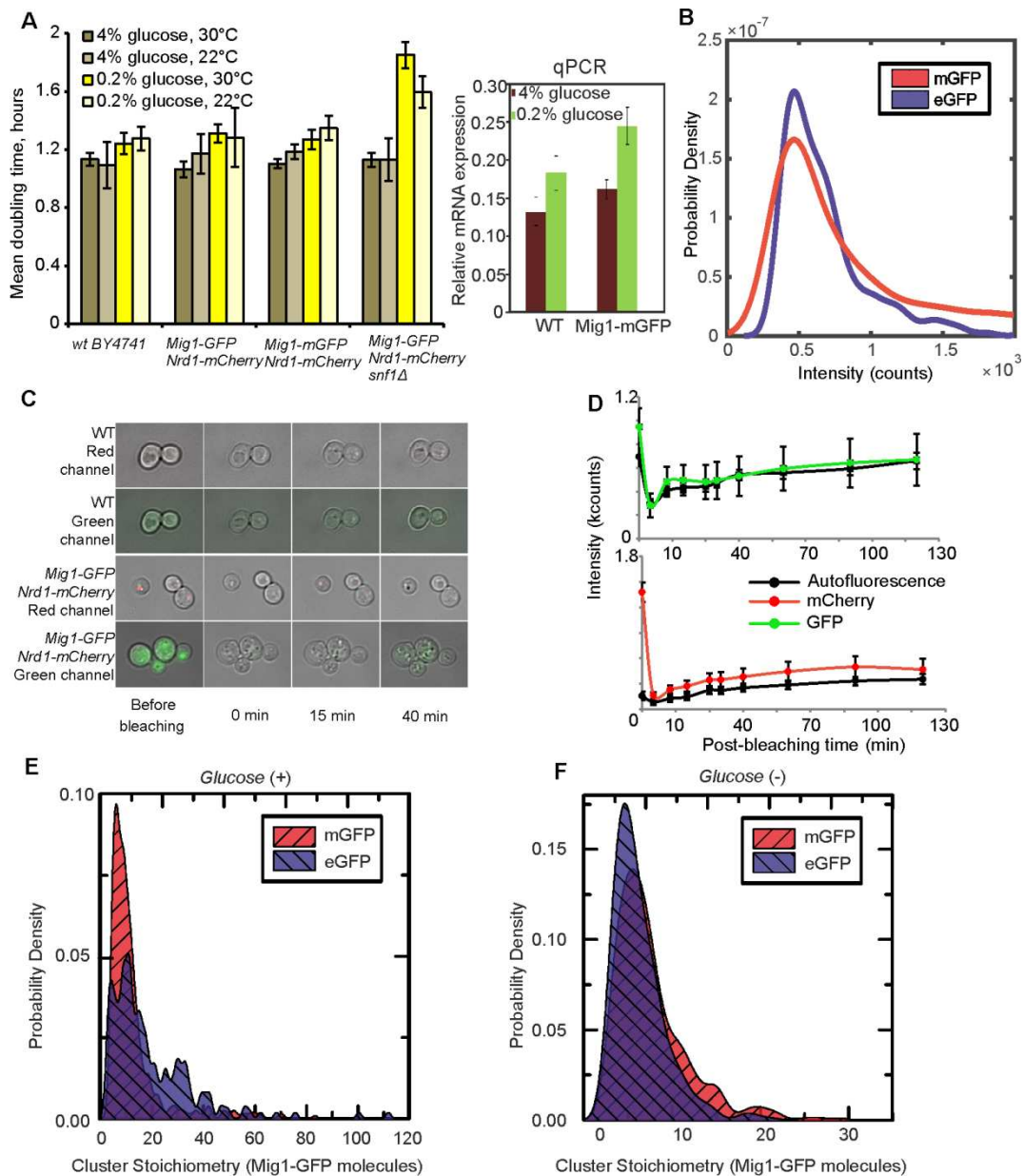


Figure 1 – Figure Supplement 2. Fluorescent reporter strains have similar viability to wild type, with relatively fast maturation of fluorescent protein, and no evidence for GFP-mediated oligomerization.

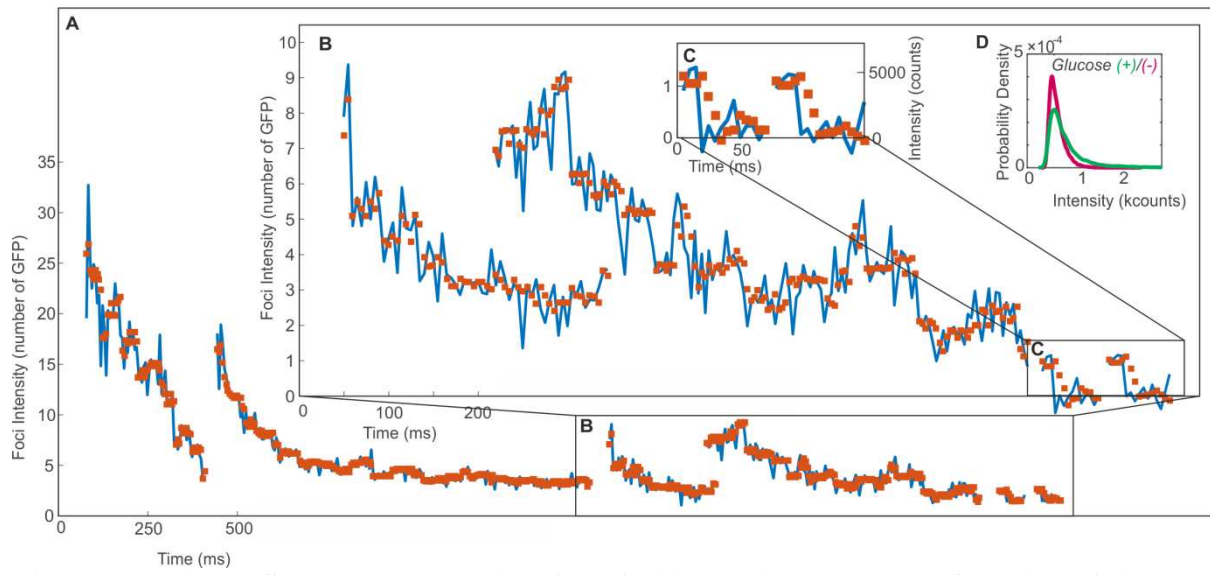


Figure. 1 - Figure Supplement 3. Mig1-GFP foci intensity traces as a function of time.

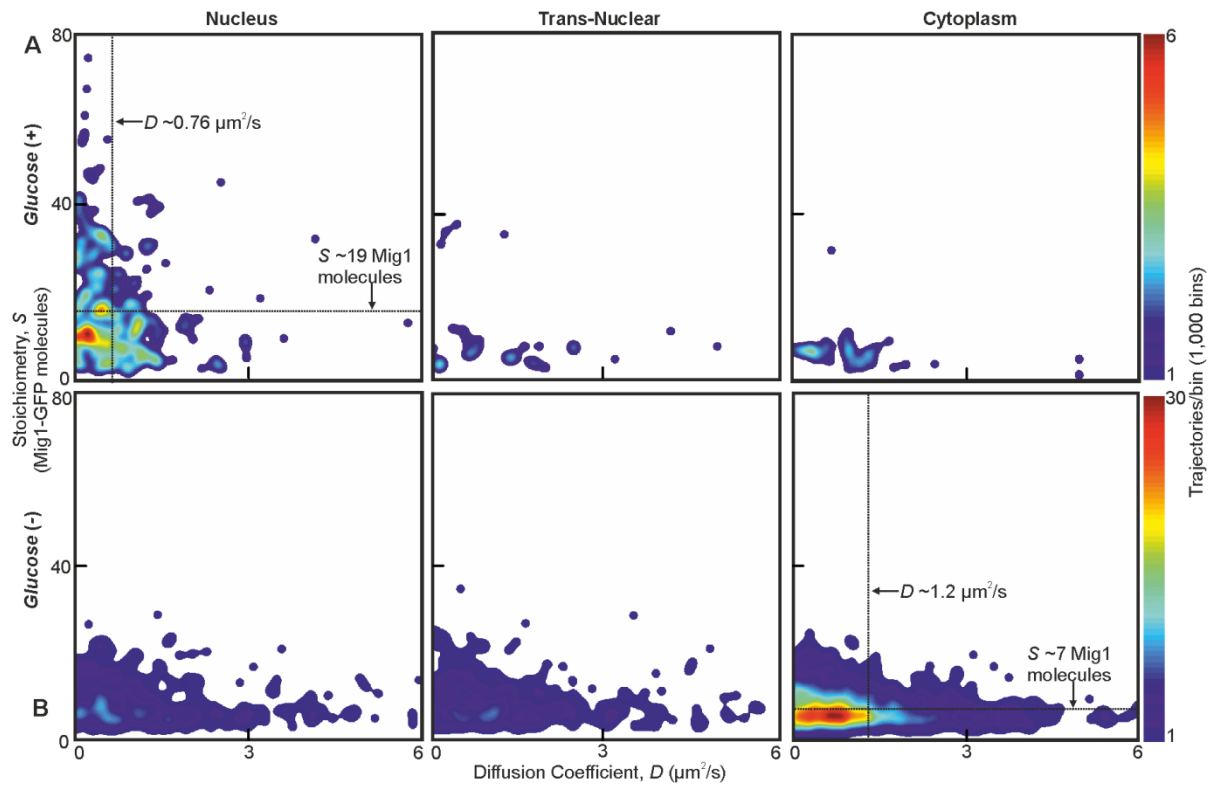


Figure 2. Mig1 foci stoichiometry, mobility and localization depend on glucose.

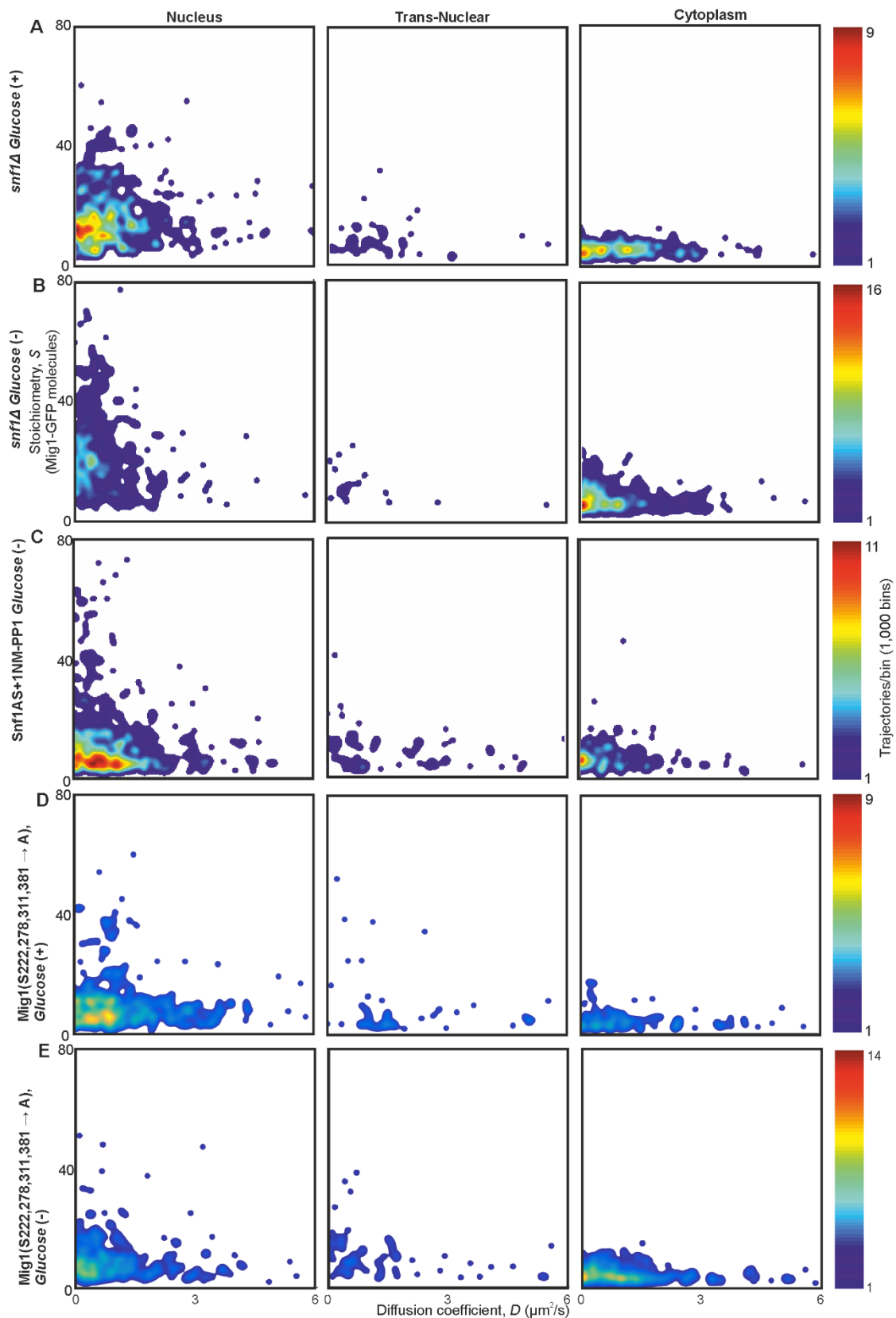


Figure. 2 – Figure Supplement 1. Mig1 phosphorylation does not affect clustering but regulates localization.

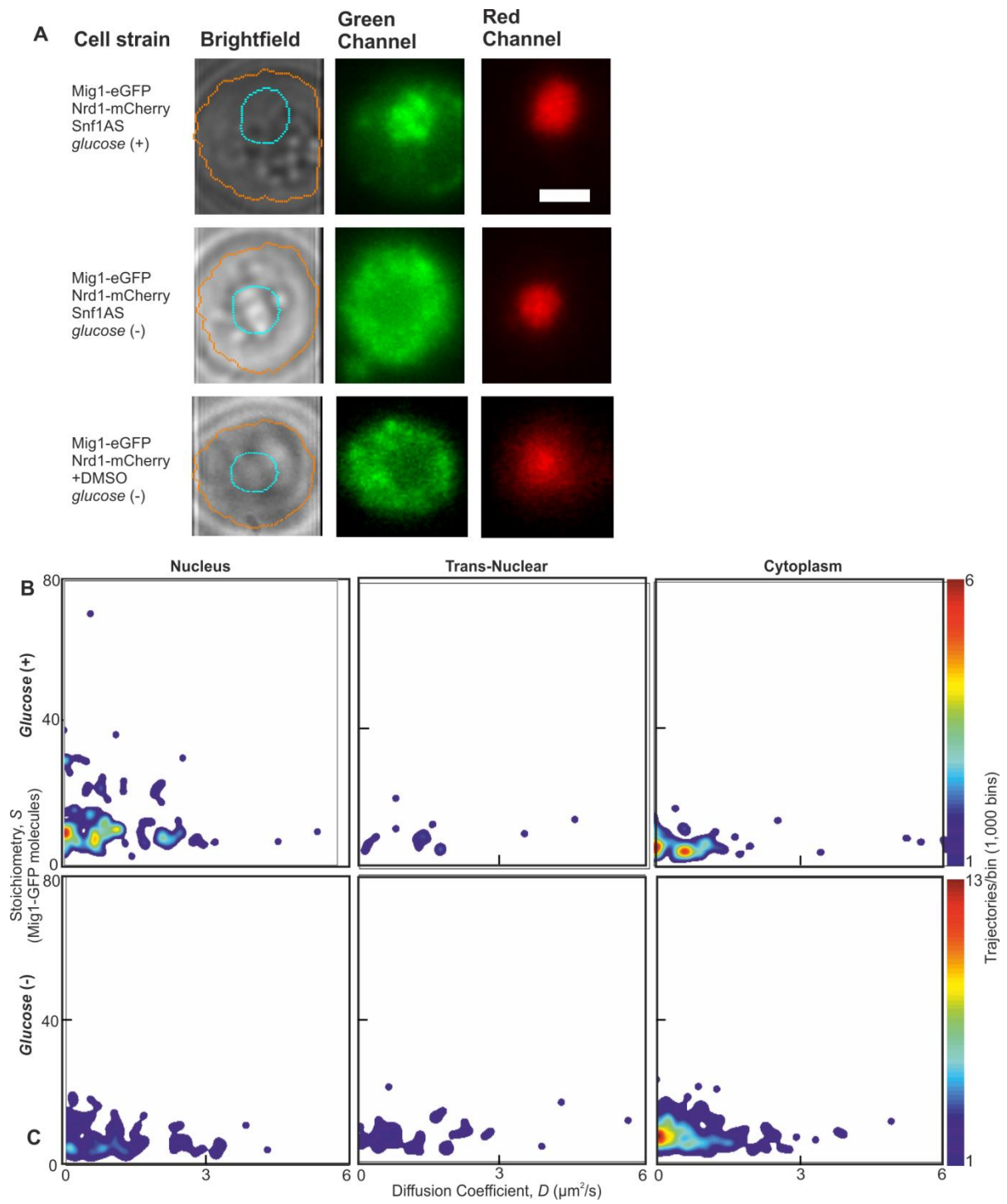


Figure. 2 – Figure Supplement 2. Wild type Snf1 and analog sensitive have similar effects on Mig1.

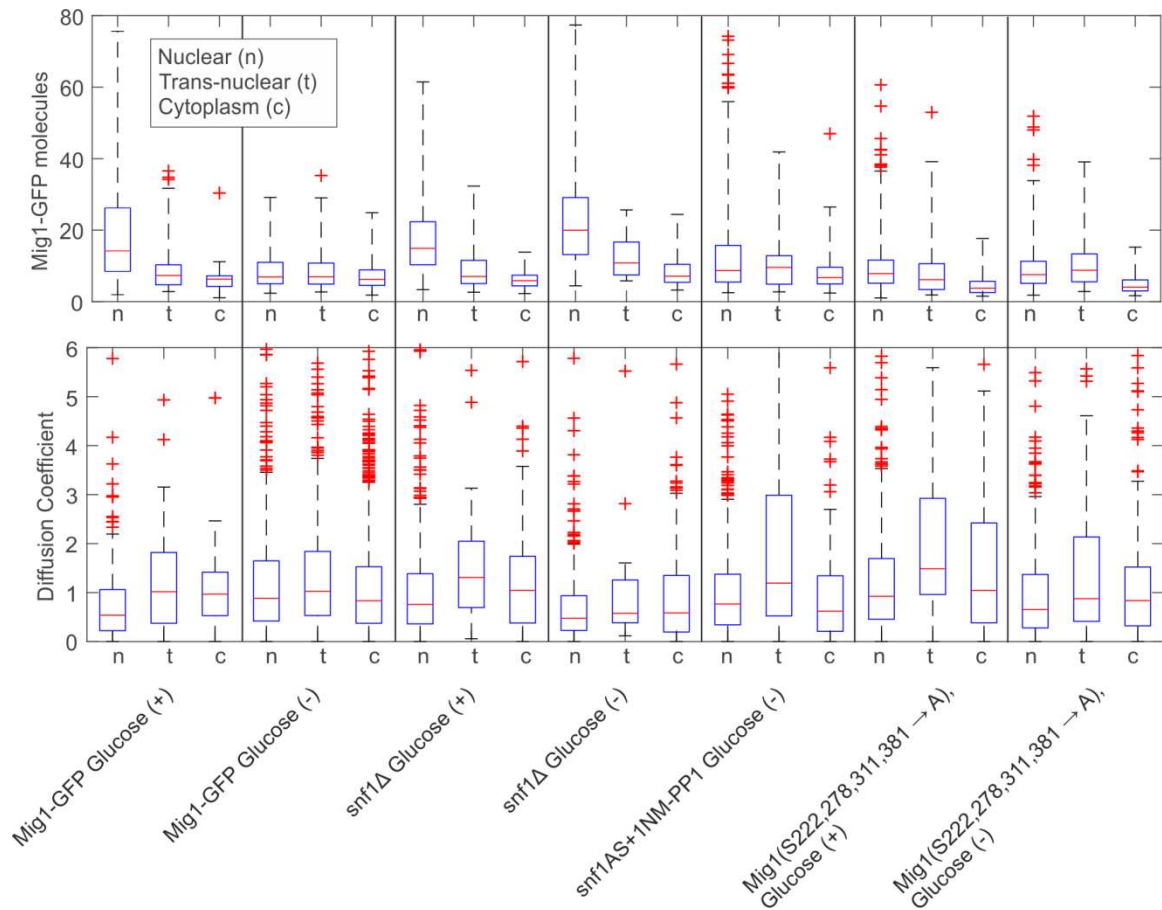


Figure. 2 – Figure Supplement 3. Boxplot summary of wild type and mutant Mig1 stoichiometry and microscopic diffusion coefficient.

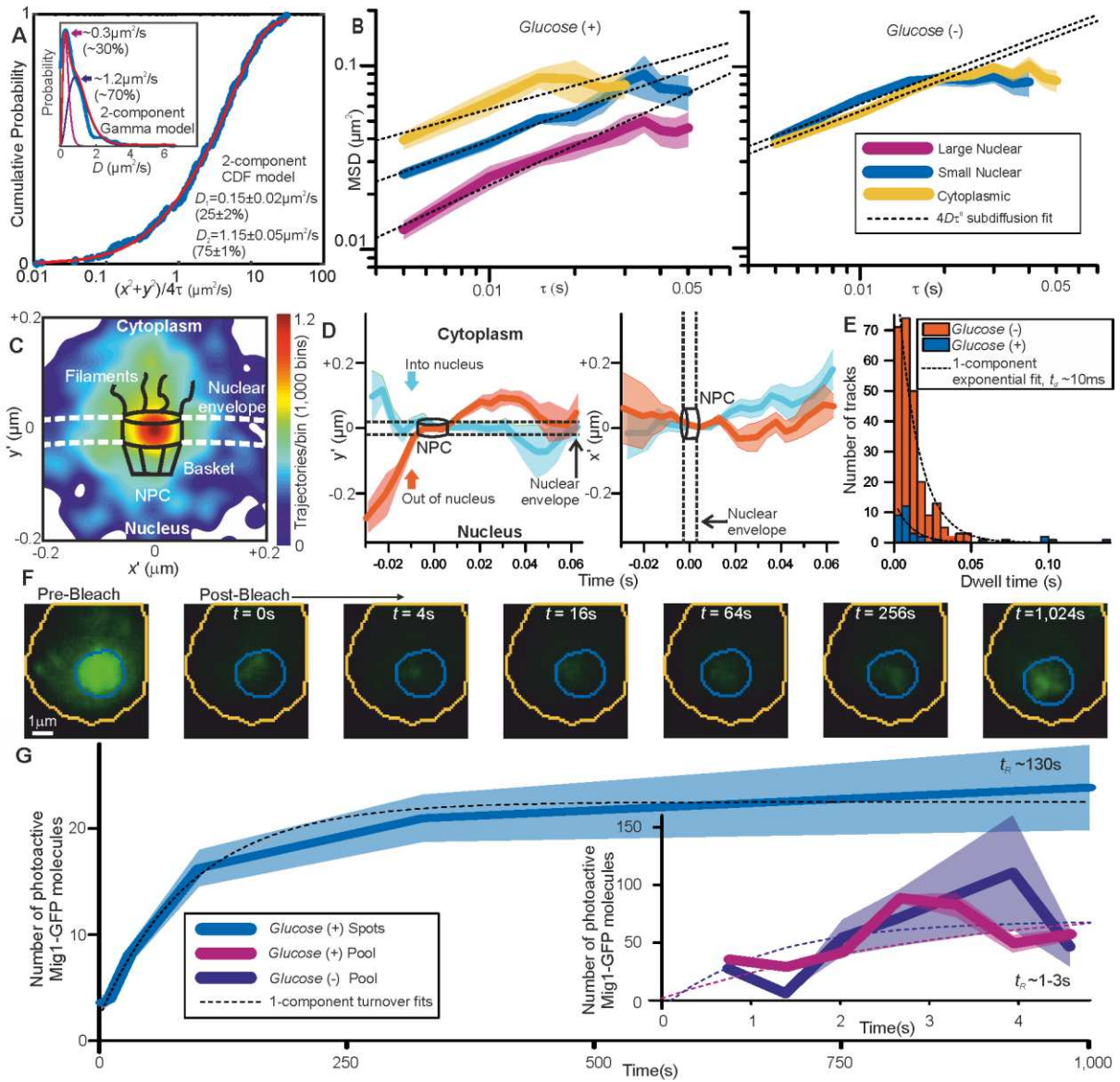


Figure 3. Repressor clusters have heterogeneous mobility depending on localization.

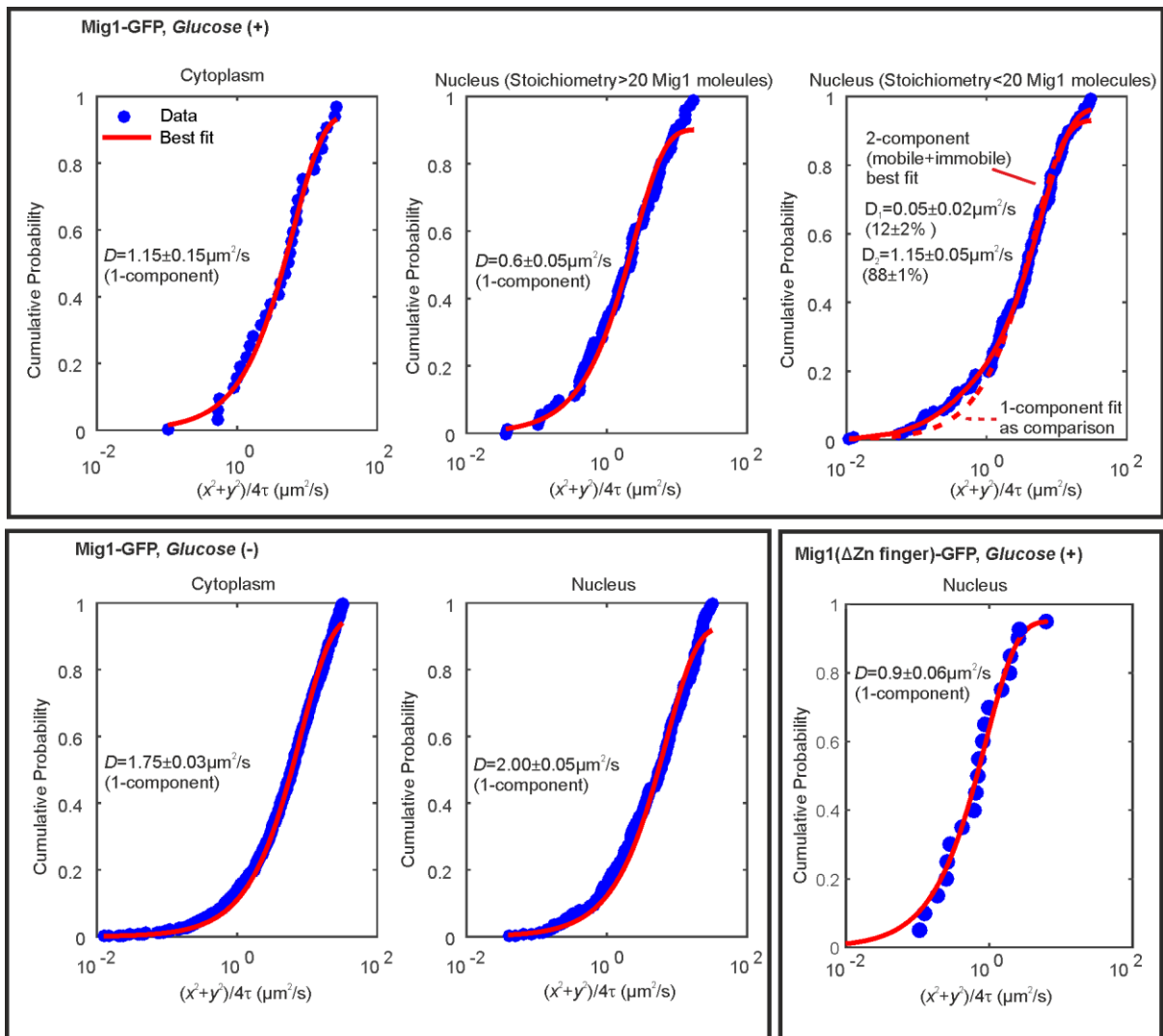


Figure. 3 – Figure Supplement 1. Cumulative probability distance analysis reveals a single mobile population in the cytoplasm at *glucose* (+/-) and in the nucleus at *glucose* (-).

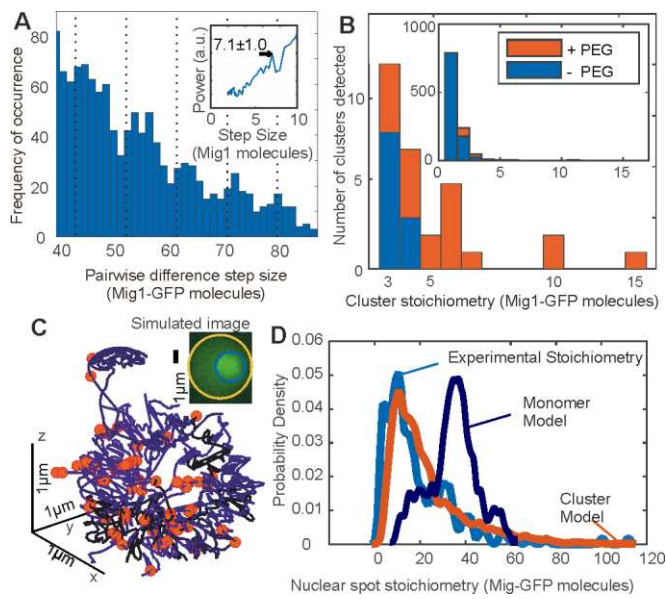


Figure 4. Mig1 clusters are stabilized by depletion forces and bind to promoter targets.

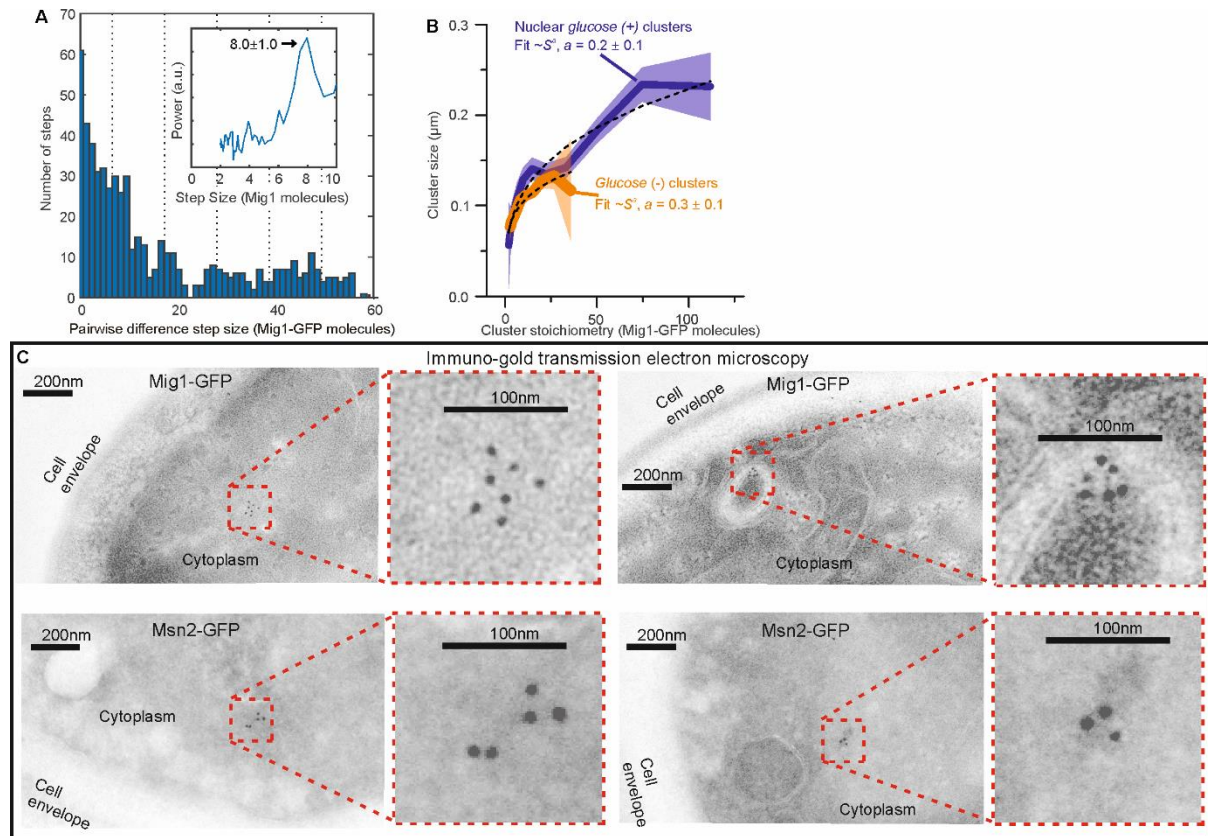


Figure. 4 – Figure Supplement 1. Additional Mig1 cluster investigations.

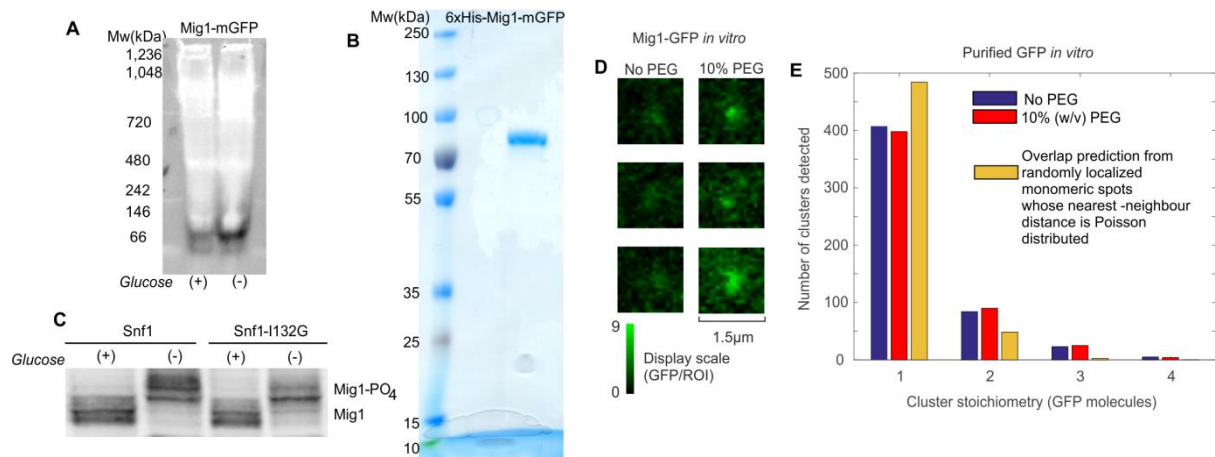


Figure. 4 – Figure Supplement 2. *In vitro* cluster characterization.

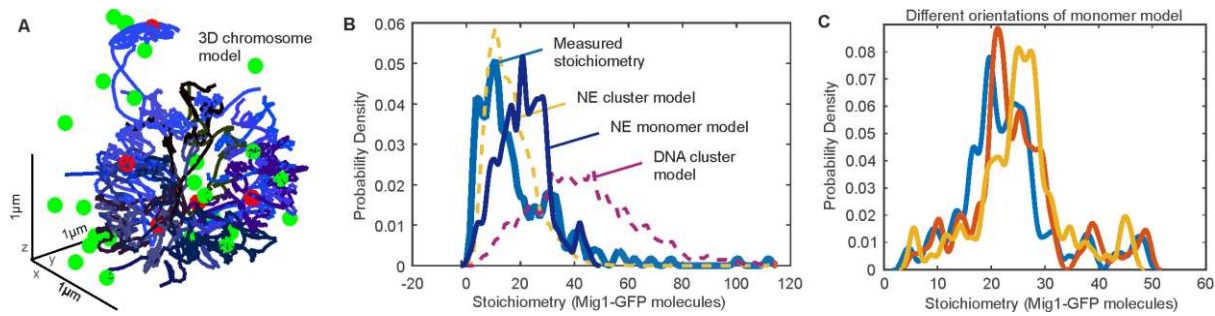


Figure. 4 – Figure Supplement 3. Additional 3C modelling.

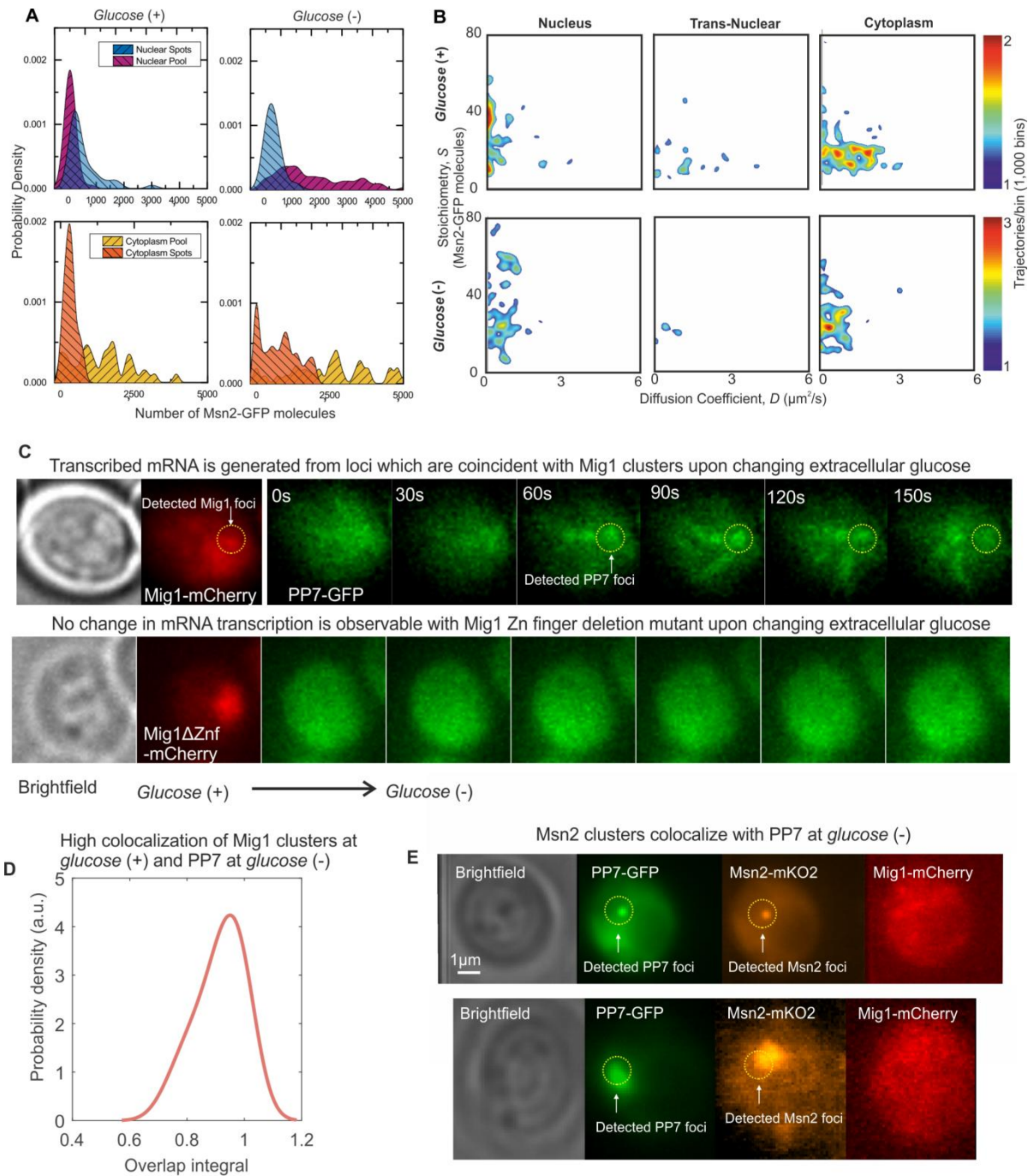


Figure 5. Msn2 and Mig1 forms functional clusters colocalized to transcribed mRNA from their target genes

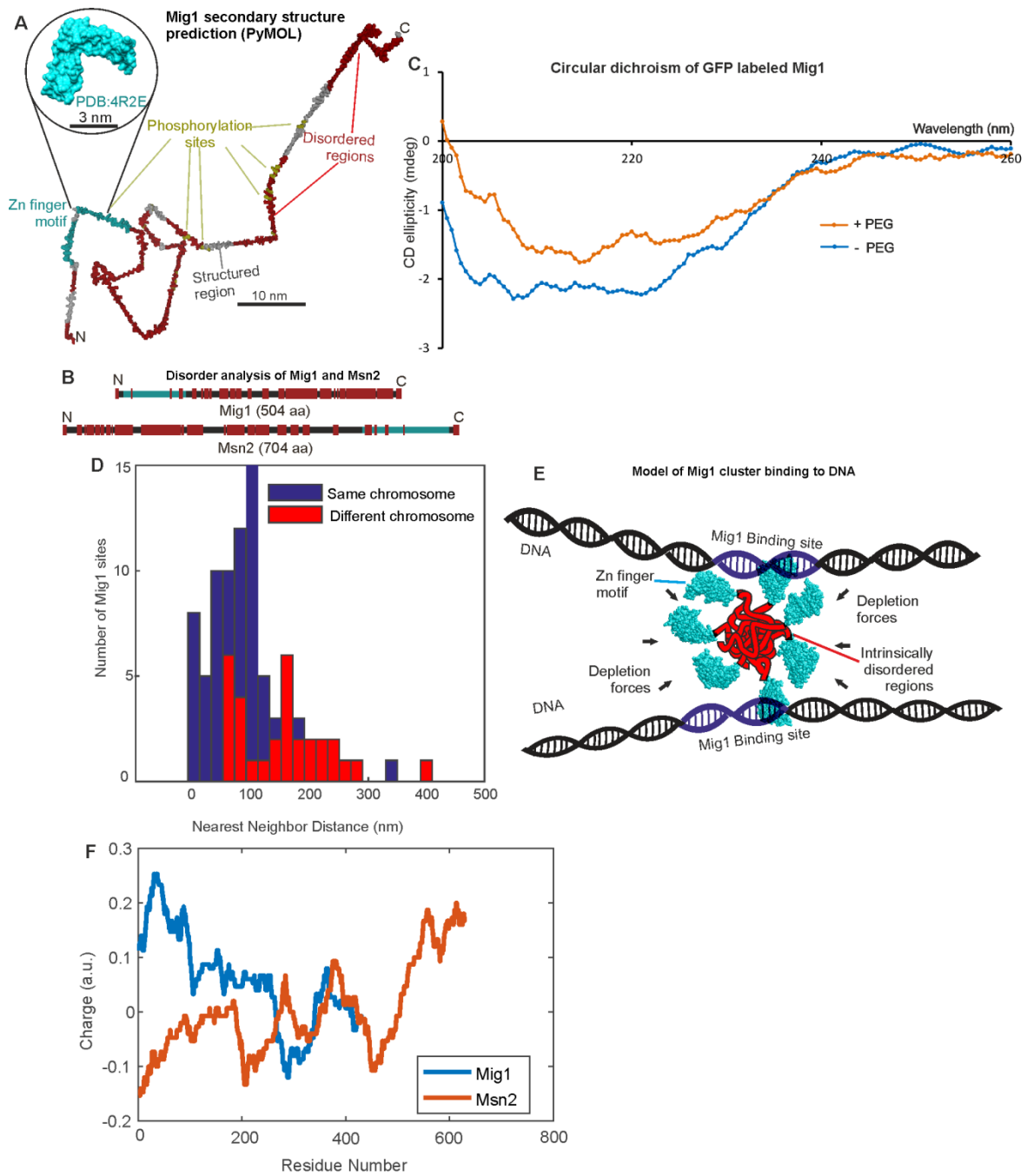


Figure 6. Mig1 and Msn2 contain disordered sequences which may mediate cluster formation.