modERN call

White Lab

2016-07-28

Approximate Status

	FY1 8/2013- 7/2014	FY2 7/2014- 7/2015	FY3 7/2015- 7/2016	FY4 7/2016- 7/2017
BACs Tagged and sent	80/80	80/80	80/80	142^/60
Transgenic Fly lines	80/80	80/80	16*/80	60
Fly ChIP-seq	60/60	70/70	80/80	2!/90
Worm ChIP-seq	60/60	70/70	30!/80	90

*: As of 4/19/16 (not including 34 new lines).
^: As of 4/6/16.
!: As of 7/25/16.

post modERN

- 322 TFs will remain after project.
  - 129 TFs left to tag
  - 83 recalcitrant to recombineering
  - 50 large BACs not sent for injection
  - 60 BACs will not have been injected

- CRISPR or reattempt recombineering
  - Most failures could be fixed by manual intervention
  - Tried CRISPR with O'Connor-Giles lab @ UofWI-Madison
    - tried EcR, yki, Eip75B.
    - All failed miserably.

BDSC

- 315 lines available

- 268 lines ordered 4797x

- Most ordered lines:
  - foxo: 182; Mad: 143; Stat92E: 134; N: 108; Dad: 106
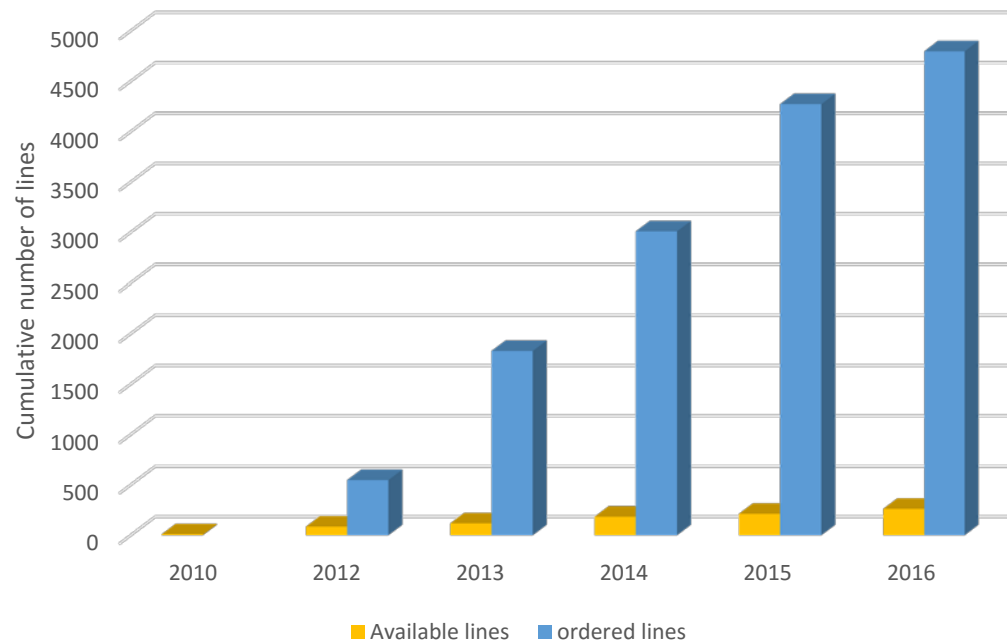
New Lines From Berkley

- 6/6 (20 lines)
  - bigmax, C15, caup, CG6791, CG7271, CG7691, CG8281, CG8301, CG9797, D, D1, E(spl)m5, E(spl)m8 , ERR, Fer2, Fer3, lab, pfk, sqz, trem

- 6/26 (14 lines)
  - phol, CG2199, CG2678, CG5180, CG7056, CG8216, CG13296, CG13894, CG14711, CG14962, CG15514, CG15696, CG15710, CG17802

- 36 more lines need to be sent to BDSC

modERN ChIP-seq datasets

228 lines complete data sets

- ab, **Abd-B**, achi, acj6[!], ADD1-mimic, Antp-mimic, Atf-2, Atf3, az2, Bab2, br, brk, Bteb2, btn, **cad**, Camta-mimic, cato, CG10462, CG10565, CG10631, CG10654, CG11398, CG11762, CG11902, CG12104, CG12155, CG12236, CG12744, CG12942, CG13123, CG13624, CG14965, CG15073, CG1602, CG1620, CG1647, CG16863, CG1792, CG1832, CG18476, CG18764, CG2116, CG2120, CG30403, CG3065, CG31388, CG32264, CG31627, CG3163, CG33213, CG3838, CG3995, CG4282, CG4402(aka CG34406), CG4424, CG4617, CG4820, CG4854, CG5204, CG5245, CG6792(aka Plzf), CG7368, CG7786, CG8089, CG8319, CG8944, CG9139, CG9305, CG9609, CG9876, CHES-1-like, chif*, chn, Chrac-16, **cic**, cnc, corto, crebA-mimic, crc-mimic, crg-1, **crp**, cyc, da, **dac**, Dad, **Dfd**, Dif, Dip3, disco, dl*, dm, dpn(&mimic), dsf, dsx, E(bx), **EcR**, Eip75B-MiMIC, Eip78C, Eip93F, Elba2, emc, ems, en, ERR, esn, E(spl)m3, E(spl)mγ-HLH, Ets21C, Ets65A(ets3), Ets97D, E(var)3-9, **eve**, **exd**, ey, eyg, E(z), FoxP, foxo(&MiMIC), ftz-f1, fru-mimic, fu2, GATAd, gcm2, gfzf, grh, **grn**, gro, **h**, her, HLH54F, HLHm7, HmgD, Hmx-MiMIC, Hnf4, Hr38, Hr39, **Hr4**[!], Hr46, Hr51, **Hr78**, Hr83, Hr96, hsf, ind, insv, jim, **jing**, Jra*, jumu-mimic, kay, kn, lbe, lilli , **lola***, Lpt, luna-mimic, lz, Mad, maf-s, mam-mimic, Max, med, Mes4, Met, Mio, Mnt, mod(mdg4), myb, **N**, NC2alpha, NC2beta, net Neu2, NK7.1, **nmo**, OdsH[!], ovo, p53-mimic, pb, pdm3, **pdp1**(&mimic), pho, Pif1A, Pif1B, pnt-MiMIC, psq[!], pum-mimic, Rel, repo, rgr, sage, salr, sens, shn, side-mimic, sima, six4, slou, slp2, Smox, Sox102F, Sox14, Sox15, Sry-delta, **Stat92E**, Su(H), Su(var)3-7, sv, svp, tai, tin, **tio**, **tll**, toe, topi, trh, trl-mimic, tup, tx, **usp**, vfl, vri, Vsx2, woc, Xbp1, Zfh2, ZnT49B.

- **14 already submitted by modENCODE**

- 214 for modERN

* multiple isoforms run

! Multiple time-points collected

XX: verifying data with rerun.

Current ChIP-seq

MiMIC being expanded for ChIP-seq (CG16779, CG9727, snoo, dsx[†])

83 tagged lines being expanded for ChIP-seq
Target Stages:

Embryo: ac, ash1, bcd, Beaf-32, BtbVII, CG10274, CG10543, CG11723, CG13775, CG15011, CG15479(Mabi), CG15601, CG16629, CG16815, CG17181, CG17568, CG18011, CG18619(REPTOR), CG30431, CG3281, CG6654, CG6683, CG6765, CG7556, CG7928, CG7987, CG8359, CG8388, CG9817, CG9948, CrebB, CTCF, D19B, dalao, E(spl)m-beta, fd96Cb, Fer1, fkh, Hand, Hey, HmgZ, Kr, l(3)neo38, meics, mor, MTF-1, nau, Nf1, oc, odd, org-1, pita, run, sc, su(Hw), su(var)2-10-RH, TFAM, unpg, Usf, Vsx1, YL-1.

W3L: CG7839, dmrt93B, HLH106.

WPP: bsh, caup, CG17803, CG8145, CG32006, CG4318, CG6276, CG6808, CG6854, CG9139(Rabex-5), Sp1.

AM: CG11617, CG33017, CG7045, Clk.

AF: CG30403, CG8159, pad, tj.

Repeat: Abd-B, cic, dac, insv, jing, nmo-small, tio.

22: expanding
07: Chromatin extracted
25: IPed
29: Libraried
    21: @ FGF
    00: Processing
    09: awaiting reps or
        recollecting

Control ChIP-seq

- nls-GFP Controls 5629/30:   collections
    - E0-24: X X X X   Need: 0 more
    - W3L: X X   Need: 1 more
    - WPP: X X X   Need: <1 more
    - AF: X X X   Need: 1 more Ad, one failed
- W1118 goatV:
    - E0-24: X X X   Need: 0 more
    - W3L: X X   Need: 1 more
    - WPP: X   Need: 2 more
    - Ad: X X X   Need: 0 more
- TF Abs
    - cnc @ E0-24
    - eve @ E0-24
    - kn @ E0-24
    - Med @ E0-24
    - Slp2 @ E0-24
    - psq @ E0-24, input failed initial BA
    - CG1832 @ E0-24
    - Xbp1 @ E0-24
- Other Ab
    - Ab290 @ E0-24
    - FLAG @ E0-24   Still needs IP-WB validation

XX: Chromatin extracted
XX: IPed
XX: Libraried
XX: Finished

Over expressed nls-GFP

	Biorep1					Biorep2				Biorep1 unique	Biorep2 unique	overlap
	Rep1	Rep2	Rep3	Rep4	INPUT	Rep1	Rep2	Rep3	INPUT			
Em	20151022									9000		
	2015-2688	2015-2689	2015-2690	-	2015-2691	2016-1558	2016-1559	2016-1560	2016-1561			
W3L	20150805					20160411				1166	2302	4624
	2015-1886	2015-1887	2015-1888	-	2015-1889	2016-442	2016-443	2016-444	2016-445			~70%
WPP	20151022					20160520				2240	1213	4806
	2015-2567	2015-2568	2015-2569	2015-2570	2015-2571	2016-648	2016-649	2016-650	2016-651			
Ad	20151206									5724		
	2015-3001	2015-3002	2015-3003	-	2015-3004							

(33) Worm

- (8) Just finished processing:
  - TBX-9_yAd, SOX-4_LE, T26A5.8_yAd, TBX-2_L3*, Y22D7AL.16_LE, GMEB-2_LE, CEH-48_LE, CEH-36_EE
- (7) Just submitted extra reps to complete sets:
  - SOX-4_L1, LIR-3_L4, HND-1_yAD, LSY-12_L3, MLS-2_EM, B0310.2_EE, MADF-2_LE
- (12) Libraries built, awaiting sequencing:
  - TRA-4_EE, SEA-2_yAd, ZTF-3_yAd, ZIM-3_yAd, MEL-28_yAd, CEH-9_LE, ATTF-2_EE, BED-3_EM, FKH-8_L1, CEH-24_LE, BAR-1_LE, HMBX-1_EE
- (6) Need new reps:
  - F13C5.2_EE, CEH-79_yAd, CEBP-1_yAd, CCCH-3_yAd, B0261.1_yAd, ATF-1_yAD
- (-) Just received:

*needs additional input sequenced

DCC

- Fly:
  - 182 datasets released
  - 14 additional datasets submitted
    - Some awaiting final antibody info
    - Some just need manual intervention (eg psq)
  - 15 more need to be uploaded

- Worm:
  - 92 datasets released
  - 49 additional datasets submitted
    - Some awaiting final antibody info.
  - 12 need to be uploaded:
    - UNC-120_OP659, HND-1_OP573, POP-1_OP560, FKH-3_OP594, F52B5.7_OP606, CHD-7_OP609, T26A5.8_OP598, TBX-9_OP636, Y22D7AL.16_OP615, GMEB-2_OP622, CEH-48_OP631, CEH-36_OP620

- Next: upload reprocessed modENCODE data.
  - Need to finish processing old data with new genome assemblies

July DCC Call

- Update on project graphic?
- Set audits for:
  - Library Complexity and IDR same as ENCODE
  - Useable fragments is ~~5M~~ 3M
  - Generating test instance to check if any datasets are flagged with these standards.
- Current Project Description OK?:
  - The model organism encyclopedia of regulatory networks (modERN) project has merged the C. elegans and Drosophila groups focused on identification of additional transcription factor binding sites in each organism. The goal of the project is to comprehensively map the binding sites for as many factors as possible. This project is still ongoing.