Review

UHN Microarray Centre

Optimisation of ChIP-chip experimental parameters

Review of: Ponzielli R, Boutros PC, Katz S, Stojanova A, Hanley AP, Khosravi F, Bros C, Jurisica I, Penn LZ. Optimization of experimental design parameters for high-throughput chromatin immunoprecipitation studies. Nucleic Acids Res 2008, 36(21):e144

The combination of chromatin immunoprecipitation followed by microarray analysis (ChIP-chip) has become a common method for studying protein-DNA interactions. ChIP-chip involves precipitating genomic DNA using protein-specific antibodies and then labelling and hybridising the DNA to microarrays. Promoter (1), CpG island (2), and whole-genome tiling arrays (3) can be used to identify gene targets that are bound by transcription factors or other proteins of interest. The optimisation of various parameters of a multi-step procedure ensures that the end result is more reproducible and reflects the genuine biology more accurately. This study by Ponzielli *et al*. is the first to provide a comprehensive evaluation of experimental ChIP-chip design parameters.

Parameters specific to ChIP-chip, such as antibody purity, amplification method for enriched DNA, and the array hybridisation control were evaluated. Two polyclonal antibodies raised against the same region of the Myc protein were compared; one purchased from a commercial vendor and the other home-made. As one might expect, the commercial antibody provided better signal-to-noise characteristics. However, the specific activity of the home-made antibody improved when it was subjected to a purification step to remove serum proteins. Sample amplification is often required as a sufficient amount of enriched DNA cannot be obtained from a single ChIP experiment. To evaluate the bias introduced by amplification protocol, DNA samples amplified by random primer (RP)-PCR, whole genome amplification (WGA), and ligation-mediated (LM)-PCR were co-hybridised with non-amplified samples from the same source. The study found that LM-PCR with low cycle number (45-cycle) and WGA were less biased amplification methods compared with RP-PCR and LM-PCR with high cycle number (60-cycle). However, LM-PCR may bias toward the amplification of smaller DNA fragments. The hybridisation of ChIP samples against a mock treated control (no antibody) and against an IgG antibody control (direct IgG) were compared. The study also evaluated using total input DNA as a common “denominator” by hybridising ChIP and IgG samples separately against total-input DNA and then comparing their intensities indirectly. Surprisingly, the results found consistency among the different controls, as the number of hits identified in each experiment was nearly identical at any given *P*-value, although the direct designs (no antibody and direct IgG) had more sensitivity than the indirect approach.

Parameters previously evaluated for gene expression studies, such as array batch variability, dye-bias and inter-day hybridisation bias, were also assessed although this study found that no more than 10% of the arrayed spots showed large variability due to these factors. These parameters were evaluated by fitting spot-wise linear models to the results of each parameter separately and plotted the Gaussian densities (essentially smoothed histograms) of the magnitudes of these effects (4). Ponzielli *et al*. concludes that dye-bias was negligible; however recommends randomising samples across dye status to account for any unanticipated dye bias. The authors also recommend using arrays from a single batch or randomising the experiments across multiple array batches. This study also found that about 10% of the spots showed inter-day variation effects. Although the exact source of this bias remains unknown, the authors speculate that the efficiencies of the labelling reaction and column purifications and post-hybridisation slide washes may have played a role, despite the fact that the experiments were carried out by the same experienced technician.

After assessing many design parameters, Ponzielli *et al*. evaluated the combined effect of the optimised parameters by conducting a ChIP-chip study using two array platforms and validating the data using real-time quantitative PCR. The binding of the c-Myc oncogenic

transcription factor in HL60 cells, a human myelogenous leukemia cell line, was combined with University Health Network Microarray Centre (UHNMAC) CGI 12K arrays. The results obtained were further validated at the UHNMAC using Agilent 2x244 promoter arrays. This comparison found that 81% (208/256) of the hits were identified on both platforms (10% FDR). From this, 50 random genomic loci were interrogated by real-time quantitative PCR and verified the extensive binding of c-Myc in HL60 cells (overall validation rate of 90%, 45/50).

This paper serves to educate researchers about the factors that must be taken into consideration when designing ChIP-chip experiments. The authors also pose some interesting questions that could be answered in future experiments. For example, by comparing monoclonal and polyclonal antibodies, it may be possible "to identify characteristics that predict which antibodies will be most sensitive or specific in a ChIP-chip study (4)". In addition, the novel technique for studying amplification bias could be used for the future development of reduced-bias amplification protocols.

References:

1. Li Z, *et al*. A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. PNAS 2003, 100:8164-8169
2. Mao DY, *et al*. Analysis of Myc bound loci identified by CpG island arrays shows that Max is essestial for Myc-dependent repression. Curr Biol 2003, 13(10):882-886
3. Takayama K, et al. Identification of novel androgen response genes in prostate cancer cells by coupling chromatin immunoprecipitation and genomic microarray analysis. Oncogene 2007, 26:4453-4463
4. Ponzielli R, *et al*. Optimization of experimental design parameters for high-throughput chromatin immunoprecipitation studies. Nucleic Acids Res 2008, 36(21):e144