

STRATEGIES TO IMPROVE THE CLONE COVERAGE IN THE HUMAN PROTEIN ATLAS PROJECT

M. Cammenberg, B. Amini, J. Fall, L. Gong, M. Hamsten, M. Salahi, A. Skölleremo and H. Eklund

KTH - Royal Institute of Technology, Sweden

malin.cammenberg@biotech.kth.se

Introduction

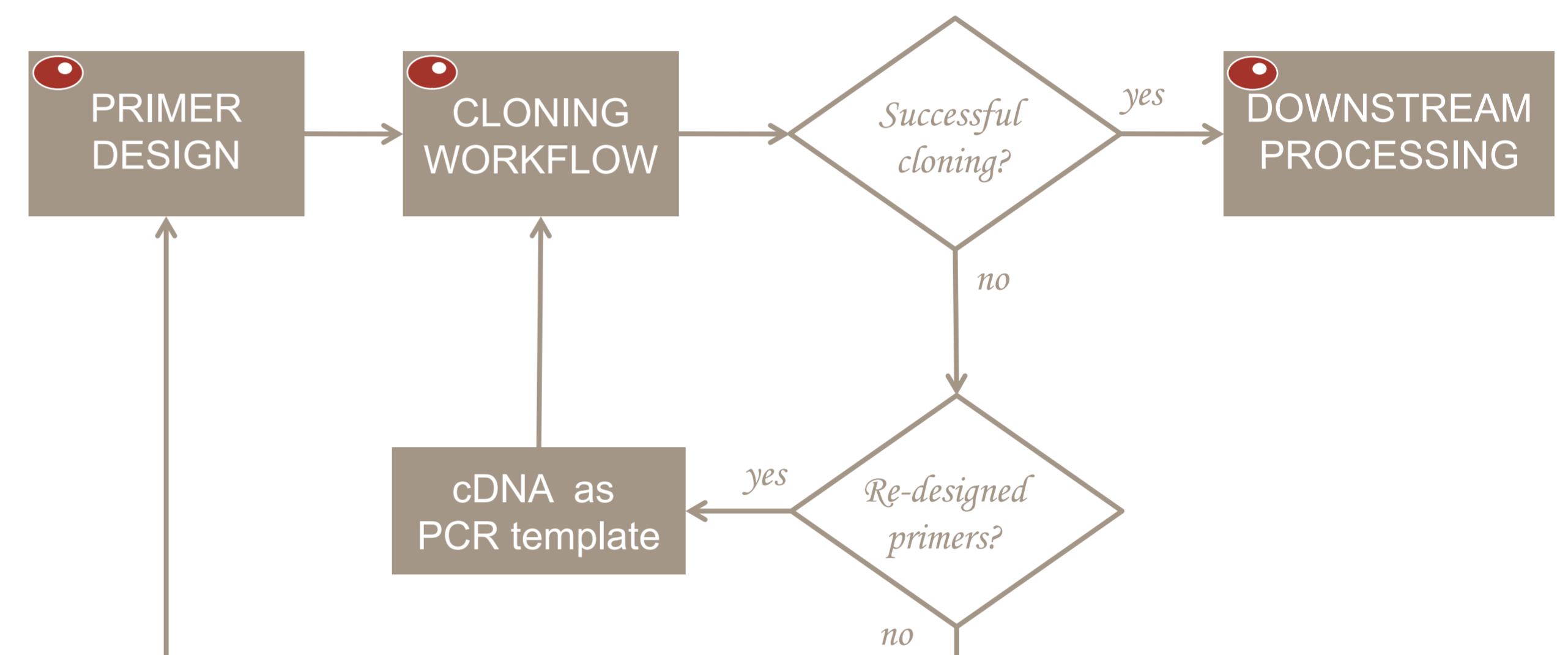
The Human Protein Atlas project (HPA) was initiated about seven years ago to characterize the human proteome using immunohistochemistry and immunofluorescence [1]. To accomplish this, unique protein regions called Protein Epitope Signature Tags (PrESTs) with low identity to other human proteins are used for cloning, protein expression and immunization to create monospecific antibodies. To date, the HPA has published over 11 000 antibodies representing almost 8 500 genes (www.proteinatlas.org).

More than 500 sequence verified PrEST clones are produced every month in the HPA. As some PrEST clones fail, new high throughput cloning strategies need to be developed continuously to improve the HPA clone coverage. Here we show results from re-design of PrESTs that have previously failed in cloning. We also describe an additional high throughput strategy using full length cDNA clones.

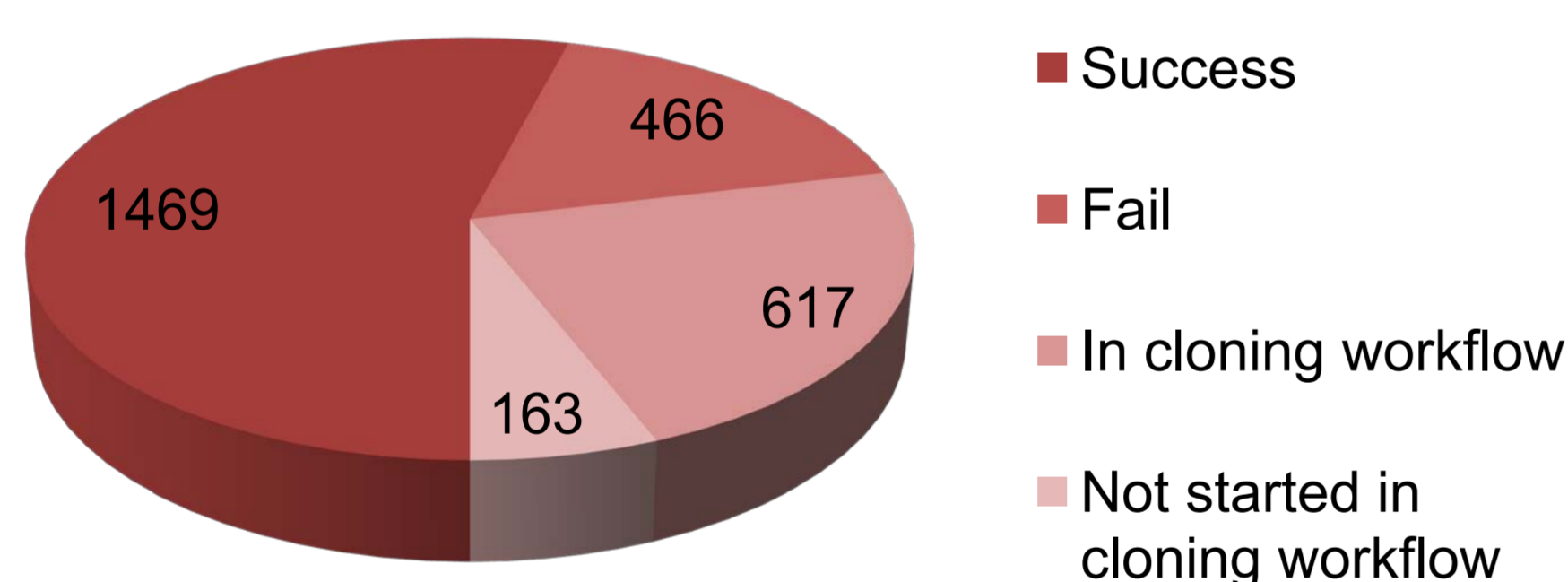
Methods

The in silico design of PrEST primers is automated by the in-house made software PRESTIGE [2]. The primers are then introduced in a standard set-up cloning workflow where total RNA from either different human tissues or cell lines is used as template in PCR amplification. If the cloning is not successful the PrEST is re-designed and run through cloning again with the re-designed primers. Although the re-design strategy is successful, some PrESTs still fail. A new strategy has been introduced where full length cDNA clones are used, instead of RNA, as template in PCR amplification.

Strategies for improving clone coverage



Primer re-design iterations



Results

To date, a total of 2 715 PrESTs have been iterated in the re-design strategy earlier described [3], and 1 469 of those have passed cloning. However, 466 re-designed PrESTs have failed cloning and are further processed using full length cDNA clones to increase the HPA clone coverage. Three different cDNA clone collections have been used; Origene, Mammalian Gene Collection (MGC) [4] and Invitrogen Ultimate ORF LITE. From Origene, 5 of 6 clones were successful. From the MGC, 13 of 32 clones have succeeded so far while the remaining 19 PrESTs are still in the cloning workflow. Of the 34 cDNA clones from Invitrogen Ultimate ORF LITE, 28 showed the correct theoretical fragment length when analyzed by gel electrophoresis, and were transferred to the cloning workflow.

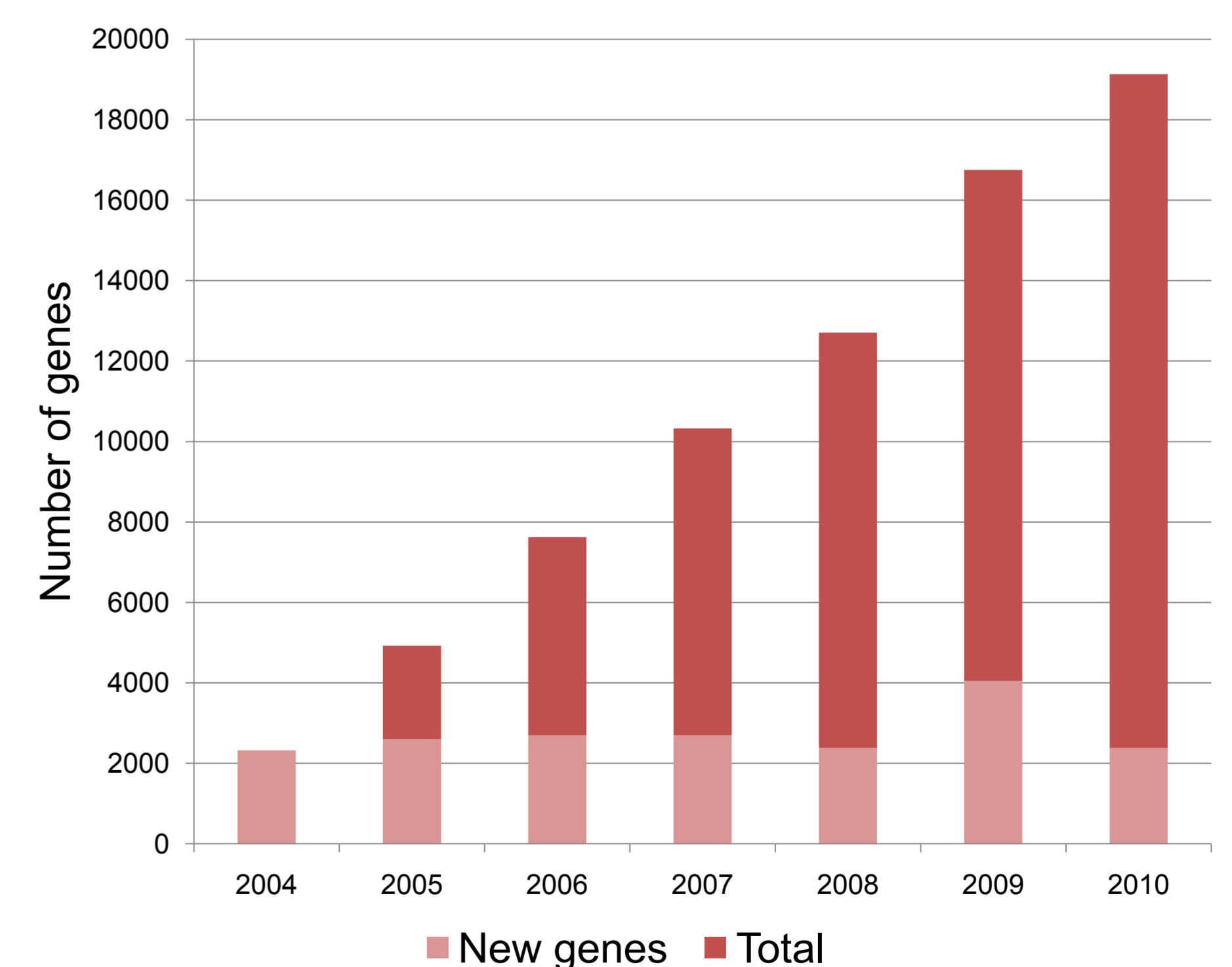
Conclusions

Both commercial companies and non-profit organizations provide a wide range of full length cDNA clones. Since the cDNA strategy seems to be successful and was easily incorporated into the existing high throughput cloning pipeline, this strategy was chosen for improving the clone coverage of the HPA project. To fully complete the HPA clone library other methods such as synthesized genes or tissue specific mRNA, other than the ones used in the project today, might be viable options.

References

- [1] M. Uhlén et al. *Mol Cell Proteomics*, 4 (12), 1920-32, (2005)
- [2] L. Berglund et al. *Proteomics*, 8 (14), 2832-9, (2008)
- [3] H. Eklund et al. *3rd EuPA Congress Clinical Proteomics*, 361-363, (2009)
- [4] MGC Project team, *Genome Res*, 19, 2324-2333, (2009)

Human genes with sequence verified PrEST clones



Acknowledgements

Funded by Knut and Alice Wallenberg foundation

*Knut och Alice
Wallenbergs
Stiftelse*

