Published in final edited form as:
Immunol Res. 2010 July ; 47(1-3): 172–178. doi:10.1007/s12026-009-8148-z.

Transcriptional target-based expression cloning of immunoregulatory molecules

Rebecca L. Lamason, **Stefanie M. Lew**, and **Joel L. Pomerantz**
Department of Biological Chemistry, Institute for Cell Engineering, and Graduate Program in Immunology, The Johns Hopkins University School of Medicine, 733 N. Broadway, BRB 607, Baltimore, MD 21205, USA

Abstract

Immunologic research has benefited tremendously from expression-cloning strategies designed to isolate genes responsible for a wide variety of immunomodulatory activities, including cytokines, receptors, signaling proteins, and transcription factors. Here, we discuss the use of expression-cloning strategies that have been modified to detect cDNAs that influence gene expression as assayed by a transcriptional reporter. We summarize our experience with these screens, review important parameters, and discuss potential modifications.

Keywords

Expression-cloning; Gene discovery; NF-κB; HIV-1; cDNA library

Introduction

Our expansive understanding of the development and function of the mammalian immune system depends upon the appreciation of the molecular machinery that governs cellular and tissue responses. Despite the wealth of understanding of innate and adaptive immunity, gaps in our collective knowledge continue to suggest that many important players are yet to be discovered.

A wide variety of approaches have been described for the identification of novel relevant molecules. These include, for example, biochemical purification of an activity followed by mass spectrometry, interaction cloning by yeast two-hybrid, RNA interference screens, and screening *in silico* for molecules with interesting homology to known regulators. Another broad approach has involved expression-cloning, in which gene libraries are expressed in a heterologous cell or tissue, an activity is assayed, and then the gene responsible for the desired activity is isolated and identified.

Our laboratory has modified the expression-cloning approach to isolate cDNAs that influence gene regulation by incorporating into the screen the quantitative assay of expression of a particular reporter gene. This strategy has allowed us to survey a wide variety of biological activities, including the components of signaling pathways that influence gene activity as well as transcription factors that act directly upon the reporter gene target. The methodology is broadly applicable, unbiased, and easily modified to enrich for genes that satisfy a particular

joel.pomerantz@jhmi.edu .

biological interest. Here, we describe our experience with this methodology and discuss its critical parameters and potential.

Cloning molecules that signal to NF-κB

We established this approach [1] with the aim of isolating novel components of signaling pathways that activated NF-κB, a pleiotropic transcription factor that is critical for both innate and adaptive immunity [2]. Pools of cDNAs from expression libraries (100 cDNAs/pool) were transiently transfected into HEK293T cells along with Igκ_2-IFN-LUC, a luciferase reporter driven by two NF-κB binding sites, and Csk-LacZ, an internal control vector for normalization that expresses β-galactosidase activity independently of NF-κB. Luciferase and β-gal activities were measured by standard techniques, and cDNA pools that activated Igκ_2-IFN-LUC by at least threefold, as compared to empty expression vector, were considered positive in the primary screen. To confirm which positive pools induced the reporter activity through NF-κB, each was tested in a secondary screen using a reporter in which the NF-κB binding sites had been mutated, MUT-IFN-LUC. Interesting positive pools were further tested in subsequent secondary screens using a dominant negative mutant of IKKβ to identify which pool activities depended on this common component of most NF-κB-inducing signaling pathways. To purify the individual cDNAs responsible for each pool's activity, pool DNA was used to retransform bacteria, and subpools of smaller complexity (24 cDNAs/pool) were cultured in 24-well plates (1 clone/well) and then screened to find the subpool that harbored the activity. Once a positive subpool was identified, the individual well containing the active clone was identified by screening pooled rows, columns, and levels from a conceptual 4 (column) × 3 (row) × 2 (level) matrix, which yielded the well coordinates of the active clone. The isolated active clone was then rescreened and sequenced.

It was initially surprising how well this approach worked. We identified a wide variety of molecules, including receptor ligands, cell-surface receptors, several kinds of intracellular signaling molecules, both catalytic and non-catalytic, as well as subunits of NF-κB itself (Table 1). The cDNAs detected exhibited a wide range of reporter activation in the assay, between 2.3- and 256-fold. The degree of activation of each molecule was not predictive of either the biochemical type of molecule or the pathway in which the molecule operated.

We then focused on identifying novel components of lymphocyte signaling pathways to NF-κB. To do so, we screened a mouse thymus cDNA library to survey genes expressed in lymphocytes, and cloned CARD11 (also known as CARMA1 or BIMP3), a multi-domain adapter protein. We and others subsequently demonstrated that CARD11 is required for TCR and BCR signaling to NF-κB and antigen-induced lymphocyte activation [1,3–8]. Following the validation of the importance of this expression-cloning isolate, CARD11 has since become a major focus of research in the laboratory [9].

Cloning a factor that can reactivate latent HIV-1

A second published example of screening was aimed at identifying cDNAs that could induce the reactivation of HIV in latently infected $CD4^+$ T cells without causing T cell activation [10]. The treatment of HIV patients with highly active anti-retroviral therapy (HAART) is effective at reducing viral loads but it does not eradicate the virus, due to the maintenance of a low level of virus in latently infected reservoirs [11,12]. The notion was that the reactivation of HIV in these reservoirs in the presence of HAART could help purge patients of latent HIV that would otherwise be resistant to therapy and would not be cleared. A factor that would not activate the T cell was desirable because non-specific or global T cell activation during HIV reactivation would likely be toxic to patients. The surrogate screen for such factors was designed to detect cDNAs that could activate transcription from the HIV-1 LTR in an NF-κB-independent manner since NF-κB activity is associated with T cell activation. We thus screened

pools of cDNAs (100 cDNAs/pool) for the ability to activate mκB-LTR-Luc, a luciferase reporter containing the HIV-1 LTR with mutated NF-κB binding sites in the core enhancer region. Our most promising clone from this screen encoded an alternatively spliced variant of the Ets-1 transcription factor, ΔVIIEts-1. We subsequently showed that overexpression of ΔVII-Ets-1 could induce the production of virus from primary $CD4^+$ T cells isolated from patients on HAART, but did not cause T cell activation. The results illustrated the potential to use unbiased screens to identify novel strategies for HIV reactivation and eradication [10].

How does overexpression lead to detection in the screen?

To conceptualize screens that are aimed at identifying signaling molecules, like the one that yielded CARD11, it is instructive to consider possible mechanisms by which the variety of molecules are detected. For example, a receptor *ligand* may be detected as an activator of a luciferase-based reporter through autocrine stimulation of the transfected assay cell line, assuming the line endogenously expresses an appropriate receptor for the ligand that could signal the activation of pathways that stimulate the reporter. Supported by our practical experience, a *receptor* may be detected in this screen whether or not a natural ligand is present in the culture assay, due to the fact that many receptors exist in inactive and active conformations that are governed by an equilibrium constant. Even if the equilibrium heavily favors the inactive conformation, overexpression may yield enough of the active conformers to activate the signaling pathway, even in the absence of a ligand that stabilizes the active state. A similar consideration can explain the detection of an intracellular *catalytic signaling protein*, which is also likely to exist in inactive and active states, the latter of which may accumulate sufficiently during overexpression to initiate downstream signaling. Overexpression of a catalytic signaling molecule may also promote its association with an endogenous modifier that could favor conversion to the active state, for example, by allosteric control. Alternatively, the overexpression of the signaling enzyme may allow access to substrates or effector molecules that cause downstream signaling by overriding cellular controls that otherwise enforce separate cellular localization of these molecules in the absence of a natural stimulus. Lastly, an overexpressed *adapter protein* may elicit activity by promoting the interaction of an endogenous catalytic molecule with its substrate or effector, by converting the catalytic molecule to an active state, or by titrating out an endogenous inhibitor of signaling.

It is straightforward to imagine how *transcription factors* are detected, since they might operate directly on the reporter being assayed or indirectly on endogenous genes whose expression leads to reporter induction. Although many transcription factors are regulated by posttranslational modification, exclusion from the nucleus, or other mechanisms, overexpression can often override these controls to produce enough activity for detection.

Considerations for screen design

This screening methodology is highly modular, consisting of a cDNA expression library, reporter construct, and assay cell line. Any of these can be substituted or modified to focus a screen on a biological target with particular properties.

The *cDNA expression library* should be derived from a cell line or tissue that contains the desired activity and should contain as many full-length cDNAs as is possible. Many commercially available libraries are available and some have been characterized with regard to complexity (i.e. how many different clones or genes are represented) and frequency of full-length clones. Expression libraries can also be constructed from a particular source using available kits or protocols [13]. In addition, whole-genome collections of cDNAs are now available from several commercial sources, so it is very possible to consider screening individual cDNAs or pools of cDNAs of known identity. A potential disadvantage of these

whole-genome collections is that they may not contain particular tissue-specific splice variants of a gene that could encode a desired activity not conferred by other splice variants.

The complexity of pools being screened is another variable that can be adjusted. In our screen for proteins that signal to NF-κB, we used a pool complexity of 100 cDNAs/pool because pilot studies with the TRAF2 protein (an E3 ubiquitin ligase) had suggested that this complexity would allow detection of molecules several fold weaker than this protein of "average" potency in the assay. Pool complexity can be as small as one, i.e. an individual cDNA, and can be as large as is allowed by the sensitivity of the reporter assay, but the complexity will determine the minimum specific activity that can be detected in the screen. In these screens, transfection efficiency is usually optimal at a given concentration of total DNA used for the transfection. Since this parameter is held constant, say at 400 ng per 10^5 cells, the higher the complexity of that 400 ng of pool DNA, the lower the concentration of each assayed cDNA in the pool. A pool complexity of 100 cDNAs/pool will allow the assay of 4 ng each for each of the 100 cDNAs in the pool. In contrast, a pool complexity of 10 cDNAs/pool would allow 40 ng of each cDNA to be assayed, while a complexity of 1 would allow 400 ng of the cDNA to be assayed. It is not predictable how many ng of a particular cDNA need to be transfected to achieve detectability in the assay, and this parameter will be highly variable from cDNA to cDNA, and screen to screen. Higher-complexity pools allow the screening of more cDNAs per series of assays and therefore one confronts a trade-off between sensitivity and throughput that each investigator must judge.

The *reporter* used for screening offers many opportunities. It may consist of one or more copies of an element recognized by a particular transcription factor, and to a first approximation, the more copies of the element, the higher the sensitivity and dynamic range of the luciferase assay. Many reporters for transcription factors are in wide use and could be used for screening. The reporter may also contain a natural gene promoter, enhancer, or other regulatory element, and this screening methodology may be used for the study of an uncharacterized regulatory element. We have used luciferase-based reporters, but other modalities should be compatible with screening as long as their outputs can be quantitatively determined and normalized for transfection efficiency and extract recovery.

The *cell line* or tissue to be used should support the biological activity that is of interest, and should be as transfectable as is necessary to detect basal and induced reporter activity. We have routinely used HEK293T cells because of their ease of transfection and because they support all activities that we have pursued. Although these cells are derived from human embryonic kidney, they can support the activity of signaling proteins that otherwise are thought to operate in a tissue-specific manner. For example, HEK293T cells could support CARD11 activity when overexpressed, even though CARD11 is not expressed endogenously in these cells. Some activities, in contrast, may require the screen to use a particular cell type or tissue and may even require the use of primary cells.

Limitations to consider

Expression-cloning is by nature a risky venture since it is difficult to predict how easily an unknown gene will emerge. Activities that are composed of more than one protein, as is the case for many protein complexes, will be resistant to identification by this methodology, given the low probability that all relevant cDNAs will be present in the same pool or assay. In addition, screening in pools entails the possibility that a pool will contain two or more active cDNAs that interfere with each other's detection. For example, if an activator of a reporter is contained in a pool that also expresses an inhibitor, the activator may remain undetected. Some cDNAs display surprising dose–response curves in reporter assays, often displaying more activity at low concentrations and less activity at higher concentrations. Such cDNAs may

elude detection if the screening assay is done at only high concentrations. Other cDNAs may be intractable if their gene products are particularly toxic to the cell line used in the screen, or if they require an endogenous factor for activity that happens to be absent from the assay cell line. All of these behaviors are difficult to predict, and should not discourage one from embarking on a screen, but it is useful to keep these examples in mind as a screen proceeds.

Other potential screen modifications

Epistasis

Screens can be easily modified to identify cDNAs with defined epistatic relationships to known genes. For example, cDNA libraries can be screened in the absence and presence of a dominant negative variant of a particular gene to identify those cDNAs that act upstream of that gene. Similarly, RNAi can be used to transiently or stably knock down a particular gene, and cDNAs can be assessed in the reporter assay to reveal which ones require the knocked-down gene for activity. In this manner, components of a signaling pathway that operate at a particular level of the pathway can be targeted. In our laboratory, we have screened for activators of the NF-κB-responsive reporter that were inhibited by coexpression of a kinase-dead variant of IKKβ. In this way, we were able to identify those cDNAs that encoded proteins that functioned upstream of the IKK complex.

Enhancers and suppressors

Another modification that we have pursued with success in our laboratory is to screen for cDNAs that modify the activity of a known protein. For example, one can include in the screening transfection an expression vector for a known protein that activates the reporter. Pools of cDNAs can then be screened for the ability to enhance or suppress the activity achieved by that protein. In unpublished work, we have successfully identified both enhancers and suppressors of a gene of interest using this methodology.

Screening of protein variants

A third variation that has produced a wealth of information in our laboratory has involved generating a library of mutants of a particular protein that shows activity in the reporter assay, and then screening this library for loss-of-function mutants in an unbiased manner. In this way, we have performed mutagenesis screens that rapidly reveal residues that are important for function.

Screening of small molecules

Finally, the methodology described also allows the screening of small molecules or other bioactive agents that could affect reporter activity or the ability of another stimulus to induce reporter activity.

Validation

Central to the success of any expression-cloning screen is the validation that must follow the identification of any interesting clone. Since these screens often rely on the overexpression of candidates, which often exceeds physiological levels by many fold, it is very possible that some genes emerge in the screen by displaying artifactual activities that are not relevant to the biological question at hand. Secondary screens and other studies are always necessary to prove that an isolate is responsible for an activity at physiological levels of expression in the appropriate cellular or tissue context.

Acknowledgments

The work referred to was supported by RSG-06-172-01-LIB from the American Cancer Society, PO1AI072677 and RO1AI078980 from the NIH, and funds from the Johns Hopkins University Institute for Cell Engineering. R.L. is supported by Ruth L. Kirschstein National Research Service Award F31AG031689. S.M.L. is supported by a Dr. Richard and Mavis Fowler and The Foundation for Advanced Research in the Medical Sciences, Inc. (FARMS) Fellowship. J.L.P is a recipient of a Kimmel Scholar Award from the Sidney Kimmel Foundation for Cancer Research and a Rita Allen Foundation Scholar.

References

1. Pomerantz JL, Denny EM, Baltimore D. CARD11 mediates factor-specific activation of NF-kappaB by the T cell receptor complex. EMBO J 2002;21:5184–94. [PubMed: 12356734]
2. Hayden MS, West AP, Ghosh S. NF-kappaB and the immune response. Oncogene 2006;25:6758–80. [PubMed: 17072327]
3. Gaide O, Favier B, Legler DF, Bonnet D, Brissoni B, Valitutti S, et al. CARMA1 is a critical lipid raft-associated regulator of TCR-induced NF-kappa B activation. Nat Immunol 2002;3:836–43. [PubMed: 12154360]
4. Wang D, You Y, Case SM, McAllister-Lucas LM, Wang L, DiStefano PS, et al. A requirement for CARMA1 in TCR-induced NF-kappa B activation. Nat Immunol 2002;3:830–5. [PubMed: 12154356]
5. Hara H, Wada T, Bakal C, Kozieradzki I, Suzuki S, Suzuki N, et al. The MAGUK family protein CARD11 is essential for lymphocyte activation. Immunity 2003;18:763–75. [PubMed: 12818158]
6. Jun JE, Wilson LE, Vinuesa CG, Lesage S, Blery M, Miosge LA, et al. Identifying the MAGUK protein Carma-1 as a central regulator of humoral immune responses and atopy by genome-wide mouse mutagenesis. Immunity 2003;18:751–62. [PubMed: 12818157]
7. Egawa T, Albrecht B, Favier B, Sunshine MJ, Mirchandani K, O'Brien W, et al. Requirement for CARMA1 in antigen receptor-induced NF-kappa B activation and lymphocyte proliferation. Curr Biol 2003;13:1252–8. [PubMed: 12867038]
8. Newton K, Dixit VM. Mice lacking the CARD of CARMA1 exhibit defective B lymphocyte development and impaired proliferation of their B and T lymphocytes. Curr Biol 2003;13:1247–51. [PubMed: 12867037]
9. McCully RR, Pomerantz JL. The protein kinase C-responsive inhibitory domain of CARD11 functions in NF-kappaB activation to regulate the association of multiple signaling cofactors that differentially depend on Bcl10 and MALT1 for association. Mol Cell Biol 2008;28:5668–86. [PubMed: 18625728]
10. Yang HC, Shen L, Siliciano RF, Pomerantz JL. Isolation of a cellular factor that can reactivate latent HIV-1 without T cell activation. Proc Natl Acad Sci USA 2009;106:6321–6. [PubMed: 19336585]
11. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 1998;338:853–60. [PubMed: 9516219]
12. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat Med 2003;9:727–8. [PubMed: 12754504]
13. Klickstein LB, Neve RL, Golemis EA, Gyuris J. Conversion of mRNA into double-stranded cDNA. Curr Protoc Mol Biol. 2008 Chapter 5: Unit5.5.

Table 1

Examples of NF-κB-inducing expression-cloning isolates

Receptors	TNFR1
	LTβR
	CD40
	TRAMP/DR3
Ligands	TRAIL
	DLK
Adapters	MyD88
	TAB2
	RIP3
	NOD1
E3 ligases	TRAF2
Kinases	NIK
	IKKε
	Snk
Kinase Substrate	MARCKS
Small GTPase	rhoB
NF-κB subunit	RelA