

# COMPOUND PROFILING

## size impact on primary screening libraries

A recent report examines the trends in utilisation of compound libraries, as well as the impact of profiling on the size of screening decks. The drivers for choosing between large or focused libraries may be different, but the ultimate goal remains the same: improve drug discovery.

Identifying a new chemical entity with the potential to become a drug must necessarily start with screening a compound library. As high throughput screening technologies have improved, libraries are now optimised with purer compounds and fewer similar analogues. Compounds are well-characterised with respect to their drug-like qualities, toxicity, fluorescence interference and structural data from studies of compound-target interactions. This information-rich compound profiling has enabled better identification of real hits. Now the question arises: can compound profiling be used to reduce the numbers of compounds needed to screen for drugs?

HighTech Business Decisions explored this question in the recent report *High Throughput Screening 2010: Effective Strategies, Innovative Technologies, and Use of Better Assays*. 85 interviews were conducted with 52 high throughput screening lab directors and 33 executives at HTS supplier companies. The HTS directors voiced strong opinions about whether primary screening libraries should be the largest possible or smaller and more focused. Those in favour of large libraries say screening cannot be limited; there is a need to cast a wide net to find an unknown drug. Whereas those in favour of small libraries say

screening too many compounds is inefficient, it is better to use structural information to prioritise and screen fewer compounds. The following two quotes are representative of the opinions about the two strategies:

“Large libraries are better, people who can’t screen large libraries say focused libraries are more important, but I think everyone should screen the largest library they can. Even Lipinsky would say the same. We don’t really know what compounds will work, so we cannot presuppose which one is best.”

*Non-commercial laboratory*

“People are getting away from screening one million to three million compounds and going to focused libraries because of costs, the need to get things quicker, and because you don’t need 50 to 100 analogues of lead series in the screen. Once you identify a particular potential lead series, you can make focused libraries around it, either doing it *in silico* or by bench experiment screening.”

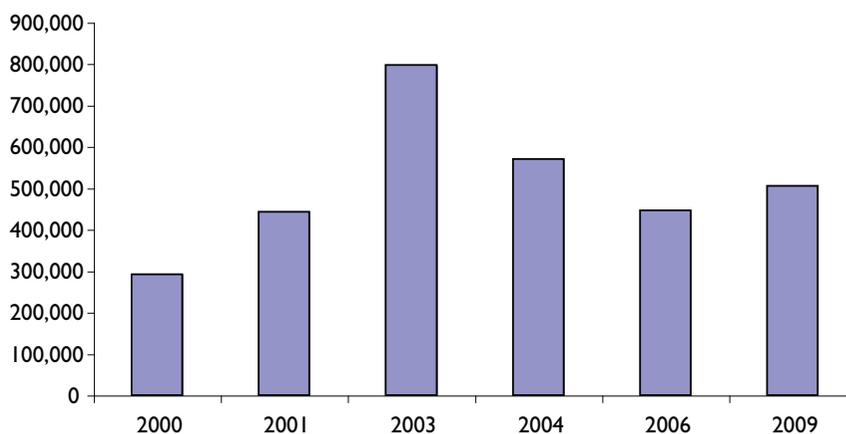
*HTS laboratory*

The current survey finds HTS labs are using an average of about half million compounds per screen for primary screening, and that they expect

**By William Downey,  
Cindy Liu and  
Dr Jennifer Hartigan**

## Compound Management

**Figure 1: Historical comparison of average compound usage for primary screens**



© HighTech Business Decisions

to increase the number of compounds they will use in 2011. Most of the HTS lab directors continue to screen the largest library they have available, although some directors report using profiling information to shift toward small, focused libraries for their primary screens.

### Compound throughput for primary screens

HTS directors report screening an average of 505,530 compounds in their primary screens in 2009. This average is consistent with the steady trend of screening of about half million compounds for primary screens observed in reports for the past five years. **Figure 1** is a historical comparison of the average number of compounds reported for primary screening from several of HighTech Business Decisions' previous reports. This comparison shows a steady increase in the number of compounds used in the early days of HTS until it reached a high of approximately 800,000 compounds in 2003. Since that high, the average number of compounds used in primary screening has settled to about 500,000 compounds. In 2011, the average number of compounds used in primary screens is expected to grow 40% to 708,123 based on opinions of laboratory directors surveyed. Most of the directors in the current study (69%) expect to increase the numbers of compounds they will use for primary screening.

The directors decide how to use their compound libraries based on many factors, including

the throughput of the assays, availability of reagents, and their knowledge and experience with the target, and consequently report using a wide range of numbers of compounds in their primary screens. The respondent labs can be divided into groups based on increasing screen size by one order of magnitude between 100 and 10 million compounds per screen. **Figure 2** shows a historical comparison of this screen size-based distribution. The most common range for screening, historically and now, is between 0.1 million and 1 million compounds, although the comparison shows a decreasing percentage of labs using screens of this size. There is a concomitant increase in the largest and smallest screen sizes, with 18% of labs now reporting screen sizes larger than one million compounds compared to 0% in 2000 and 14% of labs reporting small screens of fewer than 10,000 compounds compared with 2% in 2000.

### Opinion about the impact of profiling on selecting screening libraries

The directors agree that optimising and profiling libraries has a beneficial impact on screening, although the majority of directors report they do not use smaller, more focused screening libraries as a result of profiling. Forty-seven directors offered an opinion about large or small, focused screens as well as a variety of comments about how they optimise and prioritise their libraries and about the impact of profiling on their own screening strategies or on HTS in general. Twenty-six of these

directors said they are still running large screens despite the information they have obtained through profiling. In contrast, fourteen directors are using profiling data to run smaller, focused screens. The remaining directors said they use a mixed approach depending on the needs of their clients. A distribution of the small and large library users is shown in Figure 3; the responses are distributed by the length of time the HTS lab has been in operation. Interestingly, time of operation contributes to the strategy used by the HTS labs. Labs that have been operating for a shorter time are proportionally more likely to report shifting to smaller, focused screens.

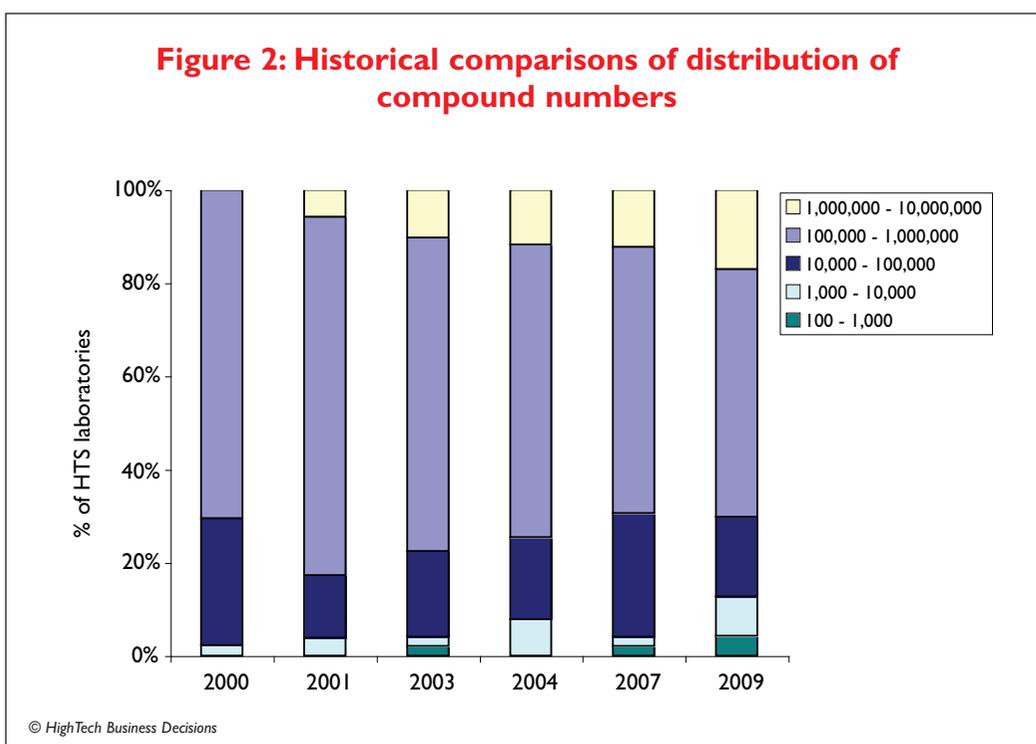
### Continue to screen large library

Labs running large screens have optimised their libraries by removing excess numbers of analogues and concomitantly increasing the diversity of the scaffold backbones in the library. The directors insist on excellent quality control of very pure compounds from the suppliers and flag interfering and toxic compounds in their current libraries. The compounds are organised with respect to their drug-like qualities and directors use this organisation as one way to prioritise compounds when they use a step-wise approach to screening the full library. Some of these optimisation approaches are described in more detail in the following comments from HTS directors:

“We use a mixture. We tried looking at subsets and found it was not as successful. We have 1.4 million compounds. If we are familiar with the target and the assay we might screen 700,000 to 800,000 compounds, but for 30% to 40% of our screens we will still use the full bank. For a new target class, we throw everything at it including the kitchen sink. We developed a new strategy where we screen a stratified library. We have clever computational chemists who have established screening clusters of 30 to 200 compounds based on tanimoto distances. We screen the library in layers, where level one includes one to two representatives of each cluster, layer two includes two different representatives of each cluster and so on. We pick the smallest and least lipophilic, the good ‘Lipinsky’ compounds first, and then we move to the more lardy compounds. We commonly stop screening at about layer seven.”

*HTS laboratory*

“Presently, our library is less than half a million, and over the next two years we will acquire more and more compounds. Our library will increase to more than two million compounds. We are looking at having more exclusive compounds that are not covered by patent or are not commercially available. Compounds in our library are better quality than they were five years ago. Our library has changed and gone through an enrichment process.



## Compound Management

For some targets we made use of a library subset when we know what we don't want to find. However, the number of full screens will not decrease, but we may have more focused screens.”

*HTS laboratory*

“Large libraries, we screen everything we have. We have reduced our library, removing a lot of the combinatorial compounds in the past several years. We have about 500,000 compounds we were screening in 2000 that we are not screening today. Our library has a greater proportion of independent scaffolds now and when we buy compounds they usually have different scaffolds.”

*HTS laboratory*

### Shift to smaller focused libraries

Directors who are shifting to smaller, focused screens are similarly prioritising their compounds with respect to the compounds' drug-like characteristics. In addition, the directors are using structural information about the target binding site and compound interactions to limit the number of compounds they are screening. The following comments by HTS directors provide detail about this structural and target-focused approach to screening:

“In big pharma, the libraries are huge, for instance three million to four million compounds and include GPCR, enzyme, kinase compounds, etc, but I believe 50% of them are not relevant to targets. So even though you do a big campaign, you don't get

that much out in the end. Our company is a central nervous system (CNS) company and while our library contains only 500,000 compounds, we believe it covers the chemical space appropriate for CNS compounds, for instance, certain molecular weight range and other characteristics. To increase their chances of getting good leads, this trend toward focused libraries is now happening in big pharma as well. They're separating their larger collections into CNS, enzyme, ion channel collections, for instance target-based libraries.”

*HTS laboratory*

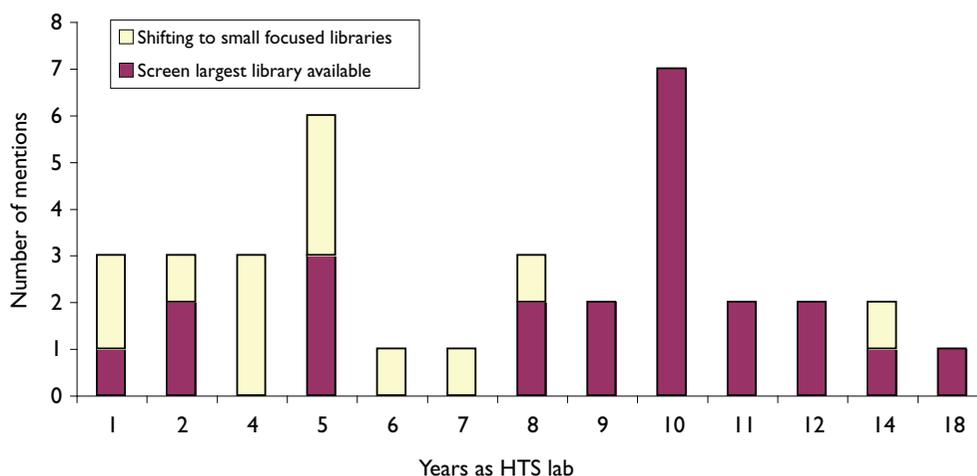
“I can only say using large libraries is not my philosophy. I don't know what other people are doing. Experience and history show that the million compound library leads nowhere, just keeps people busy. I think a focused library is what you need. The compounds should have something to do with the target and have certain characteristics, like solubility, toxicity profiles, Lipinsky rules roughly – not strictly, but enough to have an idea about the compound. And in building your library, you must start with a medicinal chemist, not a synthetic chemist!”

*HTS laboratory*

“Overall, it is going to more targeted libraries. If I could expand the library, I would expand to drugs that failed late in the game, and more targeted libraries: GPCR-ligand like compounds, nuclear receptor library.”

*Non-commercial laboratory*

**Figure 3: Large and small focused libraries: distribution by years as HTS lab**



© HighTech Business Decisions

## Compound Management

**Table 1:** Focused screening and compound prioritisation

| STRATEGIES TO IMPROVE SCREENING  | NUMBER OF MENTIONS |
|--|--------------------|
| Using the full library to screen but will use focused screens under some circumstances         | 14                 |
| Prioritised use of compounds is target or project-dependent                                    | 9                  |
| Structural information and predictive modeling are used to support and enhance screening       | 7                  |
| Step-wide use of the library is prioritised based on physical characteristics of the compounds | 4                  |

© HighTech Business Decisions

While most of the directors are not using profiling to screen smaller libraries, they all use information about the compounds to optimise their screening. The directors may have a preference for large or small library but choosing the size of the screening library is project-dependent and is influenced by factors such as previous experience with the target, physical knowledge of the target from crystal structure and fragment-based screening, or a need for a quick first look at the chemical space or limited assay reagents. A summary list about the impact of profiling information on screening is shown in Table 1.

### Conclusion

A high quality, well-understood compound library is an asset to successful drug discovery. When asked what one thing would make their screening more successful, the most frequently mentioned comment was made by 14 of the 49 HTS directors

(29%) who said they wanted expanded, diverse, high quality libraries with broader coverage of the chemical space and an increased collection of target-focused compounds.

The HTS directors have been consistently using about half a million compounds in their primary screens for several years and expect to increase the number of compounds used in the future. The directors described the optimisation and profiling they have performed on their compound libraries to improve their screening results and increase their relevant hits, but this valuable information has not prompted most of the directors to decrease the numbers of compounds used in primary screening. Although some of the directors are choosing to select and limit their compound usage based on profiling information, and this strategy is more prevalent among the younger HTS labs, most of the directors are reluctant to limit their compound usage too much in case they miss significant hits.

Compound libraries have become more refined. The compounds are higher quality and physically well defined with respect to their solubility, drug-like, toxic and interfering characteristics. However, in the end, hunting for the next as yet unknown drug means searching through a large collection of compounds. **DDW**

*William Downey is president, Cindy Liu is managing director and Dr Jennifer Hartigan is an analyst at HighTech Business Decisions, a consulting firm specialising in customised market analysis, industry reports and customer loyalty studies for suppliers serving the pharmaceutical and biotechnology industries. The company recently published the report, High Throughput Screening 2010: Effective Strategies, Innovative Technologies, and Use of Better Assays. [www.hightechdecisions.com](http://www.hightechdecisions.com)*

## ADVERTISEMENT INDEX

|                                |       |                               |       |  |         |
|--------------------------------|-------|-------------------------------|-------|--|---------|
| Abcam, Inc                     | 26    | DiscoverX Corporation         | 61    | PerkinElmer Life & Analytical Sciences | 15,66   |
| Active Motif, Inc              | 30    | EMD Chemicals, Inc            | 33    | Promega Corporation                    | 41      |
| BD Biosciences                 | 50    | Febit Holding GmbH            | 23    | Roche Applied Science                  | 36      |
| Biocius Life Sciences          | 29    | ForteBio, Inc                 | 39    | Select Biosciences Ltd                 | IBC     |
| BioTek Instruments, Inc        | 73    | HTStec Ltd                    | 17    | Society of Biomolecular Sciences       | 70      |
| BMG Labtech GmbH               | IFC-3 | ID Business Solutions Ltd     | 25    | Symyx Technologies, Inc                | 18      |
| Cambridge Healthtech Institute | 85    | Life Technologies Corporation | 13,53 | Tecan Schweiz AG                       | 8       |
| Collectis Bioresearch          | 63    | Millipore Corporation         | 47    | Thermo Fisher Scientific               | 58, OBC |
| CISBIO International SA        | 4     | Molecular Devices, Inc        | 6     | TTP LabTech Ltd                        | 42      |