cBiT

Biomaterials 149 (2017) 88–97

Contents lists available at ScienceDirect

Biomaterials

journal homepage: www.elsevier.com/locate/biomaterials

cBiT: A transcriptomics database for innovative biomaterial engineering

Dennie G.A.J. Hebels [a, *], Aurélie Carlier [a], Maarten L.J. Coonen [b], Daniël H. Theunissen [b], Jan de Boer [a]

[a] MERLN Institute for Technology-inspired Regenerative Medicine, Maastricht University, Universiteitssingel 40, 6229 ER, Maastricht, The Netherlands
[b] DataHub, Maastricht University & Maastricht UMC+, P. Debyelaan 15, 6229 HX, Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 19 July 2017
Received in revised form
20 September 2017
Accepted 2 October 2017
Available online 3 October 2017

Keywords:
Repository
Compendium
Database
Reverse engineering
Data analysis
Biomaterials
Transcriptomics

ABSTRACT

Creating biomaterials that are suited for clinical application is still hampered by a lack of understanding of the interaction between a cell and the biomaterial surface it grows on. This surface communication can strongly impact cellular behavior, which in turn affects the chances of a successful interaction between a material and the host tissue. Transcriptomics data have previously been linked to measurements of biomaterial properties in order to explain the biological mechanisms underlying these cell-biomaterial interactions. However, such multi-assay data are highly complex and therefore require careful and unambiguous characterization and storage. Failure to do so may result in loss of valuable data or erroneous data analysis. In order to start a new initiative that tackles these issues and offers a platform for innovative biomaterial development, we have created a publically accessible repository called The Compendium for Biomaterial Transcriptomics (cBiT, https://cbit.maastrichtuniversity.nl). cBiT is a data warehouse that gives users the opportunity to search through biomaterial-based transcriptomics data sets using a web interface. Data of interest can be selected and downloaded, together with associated measurements of material properties. Researchers are also invited to add their data to cBiT in order to further enhance its scientific value. We aim to make cBiT the hub for biomaterial-associated data, thereby enabling major contributions to a more efficient development of new materials with improved body integration. Here, we describe the structure of cBiT and provide a use case with clinically applied materials to demonstrate how cBiT can be used to correlate data across transcriptomics studies.

1. Introduction

The last decades have seen a surge in the design and use of novel materials for biomedical applications. While it is known that bulk material properties and the surface communication between a cell and the biomaterial can strongly impact cell shape, attachment, proliferation and differentiation, the exact interaction at the interface between cells and biomaterials is unknown and uncontrolled [1–4]. This often results in sub-optimal functioning of medical implants, increased health care costs and additional pain for the patient [5,6]. On the other hand, the knowledge gap hinders the design of beneficial bio-active surface properties.

In order to solve these problems we need to advance our knowledge by implementing an interdisciplinary method that encompasses cell biology, micro-fabrication, materials science and computational sciences. This convergence of multiple fields of research has been coined “materiomics” and represents a much needed holistic approach to the investigation of biological material systems, through the integration of biological functions and processes with traditional materials science perspectives, such as physical properties, chemical components, hierarchical structures, and mechanical behavior [2].

Although the number of studies using this materiomics strategy is still relatively limited, its frequency of use is on the rise. Especially the association of transcriptomics with material characterization data has been applied several times and more studies appear every month. Transcriptomics refers to the simultaneous measurement of all transcribed genes through either microarrays (hybridization of the transcriptome to complementary sequences of all known genes spotted on a glass slide) or RNA sequencing (RNAseq, the sequencing of the full transcriptome). In our previous

* Corresponding author.
E-mail address: d.hebels@maastrichtuniversity.nl (D.G.A.J. Hebels).

https://doi.org/10.1016/j.biomaterials.2017.10.008
0142-9612/

work, we used transcriptomics analysis and material characterization to investigate the biomaterial-cell interaction in more detail. We specifically focused on linking material properties with gene expression responses and discovered for example, a suitable model cell type for *in vitro* biomaterial testing [7], physicochemical parameters that influence differentiation [8], and possible key players in extracellular matrix deposition associated with osteoinduction [9]. Some of this work employs mesenchymal stromal cells, which are considered to be a very valuable resource for transcriptomics-based investigations of cellular processes crucial in tissue regeneration [10]. Other groups have used a similar approach, sometimes using different omics technologies such as proteomics, as reviewed by Power et al. and Gallagher et al. [11,12]. Some recent examples include a transcriptomic analysis of the effect of biomaterial substrate on the osteogenic differentiation of stem cells, an examination of early cellular and molecular responses to differently modified surfaces of medical steel, and an excellent study by Guerette et al. that integrates RNAseq with proteomics and materials science [13–15]. Materiomics has also been suggested as the way forward in oral disease diagnostics and personal health monitoring [16]. Other types of high-throughput technologies have been applied too in order to improve the efficiency of biomaterial screening while simultaneously investigating cell biology. This includes material microarrays such as the polymer-cell interaction platform developed by Anderson et al. and the TopoChip [17–19]. For an excellent overview of other high-throughput approaches for screening and analysis of cell behaviors we refer the reader to a recent review by Seo et al. [20].

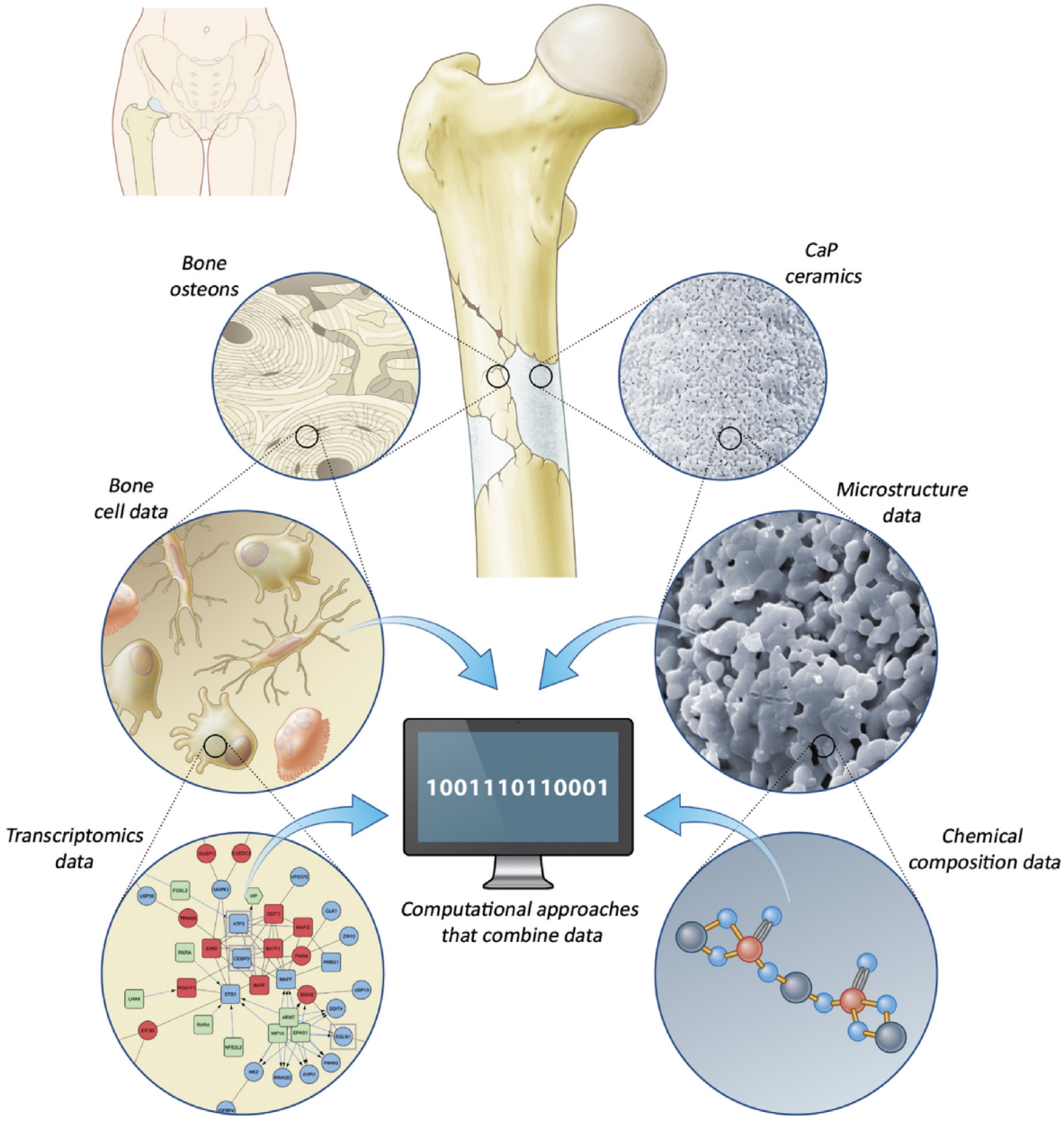

Fig. 1. Integration of biology (left) and material science (right) towards a computational approach for new biomaterial development, using bone tissue engineering of a bone fracture as an example. Both bone (left) and bone graft substitutes such as ceramics (right) can be deconstructed into their elementary components, i.e. molecular cell biology measurements and gene expression in bone cells and microstructure and chemical composition data of materials, respectively. The interplay between them can be studied using computational approaches (From Groen et al. [9]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.).

Although the first steps in the field of materiomics are very promising, these efforts have only begun to explore the complex relationship between cells and biomaterial properties. Moreover, omics technologies, as well as advanced materials engineering and characterization can generate large amounts of data on cell-biomaterial induced gene, protein or metabolite expression and material properties respectively. This is especially true for transcriptomics data and is expected to pose a major problem due to the absence of a specialized repository to collect, curate, maintain and analyze this wealth of data.

Current available repositories concentrate on either material properties (Supplementary Table S1) or transcriptomics data (Supplementary Table S2) without efficiently combining these data sets. More specifically, current materials repositories largely focus on computational material properties (e.g. Materials Project, Open Materials Database) and only a limited set of databases contain data on biomaterial properties (e.g. ASM Medical Materials Database, Biomat_dBase, Biomaterials Properties Database). Similarly, a considerable number of transcriptomics databases have been created. The Gene Expression Omnibus (GEO) and ArrayExpress are large, general purpose repositories but more dedicated databases also exist, focusing on gene expression in different species (e.g. Bgee), in cancer (e.g. Genomic Data Commons), the effect of small molecules (e.g. Connectivity Map) or toxicity (e.g. Comparative Toxicogenomics Database, diXa). While some of these repositories have the option of including metadata, such as material properties, this is not their main purpose and is therefore not done in a standardized and organized fashion, limiting an integrated analysis of data from multiple sources. Moreover, current materials repositories have a similarly narrow focus on just one piece of the puzzle, ignoring the biological questions at the root of the cell-biomaterial interface. A higher level of organization is required to solve this issue, similar to the initiatives taken by the zebrafish (https://zfin.org/) and yeast (http://www.yeastgenome.org/) research communities.

We therefore propose the establishment of a dedicated transcriptomics repository that offers an innovative strategy to improve material development while at the same time providing novel insights into the complex signaling pathways that drive the cell's response to biomaterials (Fig. 1), thereby following the principles of materiomics. In the long run, such an initiative could contribute to a more efficient development of new biomaterials that perform even better in patients. To reach this goal we present here the Compendium for Biomaterial Transcriptomics (cBiT), a publicly accessible data repository directed at collecting transcriptomics data and carefully recorded material properties and other relevant metadata (https://cbit.maastrichtuniversity.nl). A cBiT introduction video, explaining the background and goal of the repository, can be found here: https://www.youtube.com/watch?v=O12dPthanrM. Within cBiT we aim to generate and collect unique and standardized knowledge on how different cell types interact with a wide range of commonly used biomaterials. By generating new information and simultaneously accumulating it in an open access repository, we expect that it will be possible to predict cell response to biomaterials in the near future. A prediction approach towards biomaterial development can drastically reduce the current time and development costs, while at the same time ensuring higher clinical success rates. Moreover, *in silico* predictions will be essential to advance biomaterials science since the number of possible combinations of surface and chemistry properties far exceeds what could realistically be tested *in vitro* or *in vivo*.

The open access aspect of this repository will stimulate the development of new and better materials by allowing researchers to freely use data from the repository and compare their own data with repository data sets. Indeed, combining data from different studies for meta-analysis purposes has repeatedly been shown to lead to improved data interpretation [21–24]. Researchers are also invited to submit their own data to cBiT. To ensure an unambiguous characterization of data files and experiment characteristics, ISA-Tab-based data archives for each incorporated study are used [25]. There will be control on data completeness and standardization of study properties through the use of ISA tools. Ultimately, the repository-associated research could lead to the identification of genes, pathways, expression profiles or specific material properties that can inform the design and development of new implant biomaterials with excellent *in vivo* properties (Fig. 2).

2. Data infrastructure and use

cBiT comprises a central warehouse containing data from

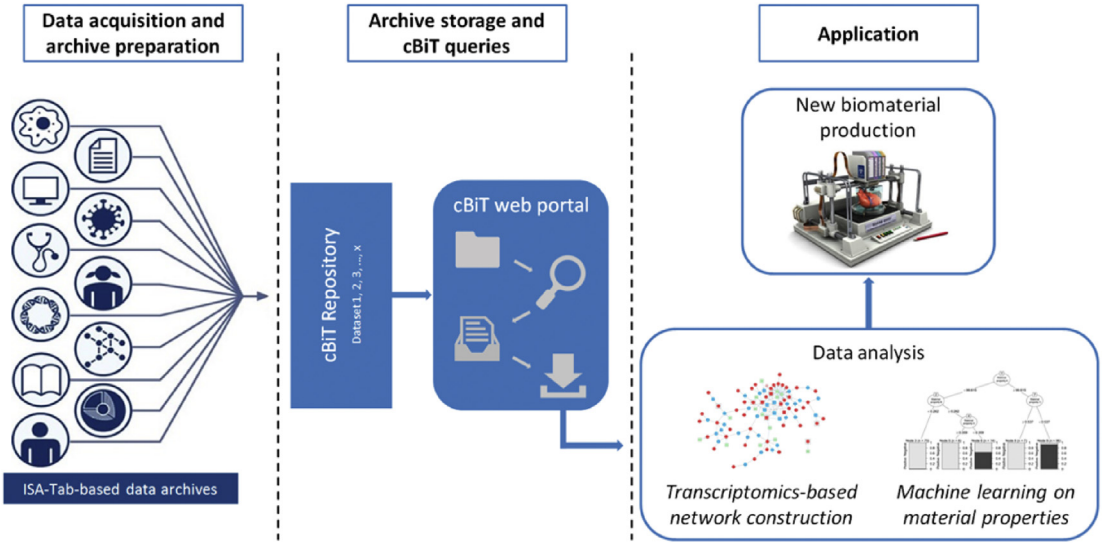

Fig. 2. Data in cBiT will be archived in a standardized way, allowing for efficient data analysis strategies. Biomaterial studies can be incorporated into cBiT using ISA-Tab archives. Once deposited, queries can be made and single or multiple data sets can be downloaded for a wide array of analysis options. This provides scientists with a powerful tool to improve their research, ultimately contributing to new cell biology knowledge and/or new biomaterial development.

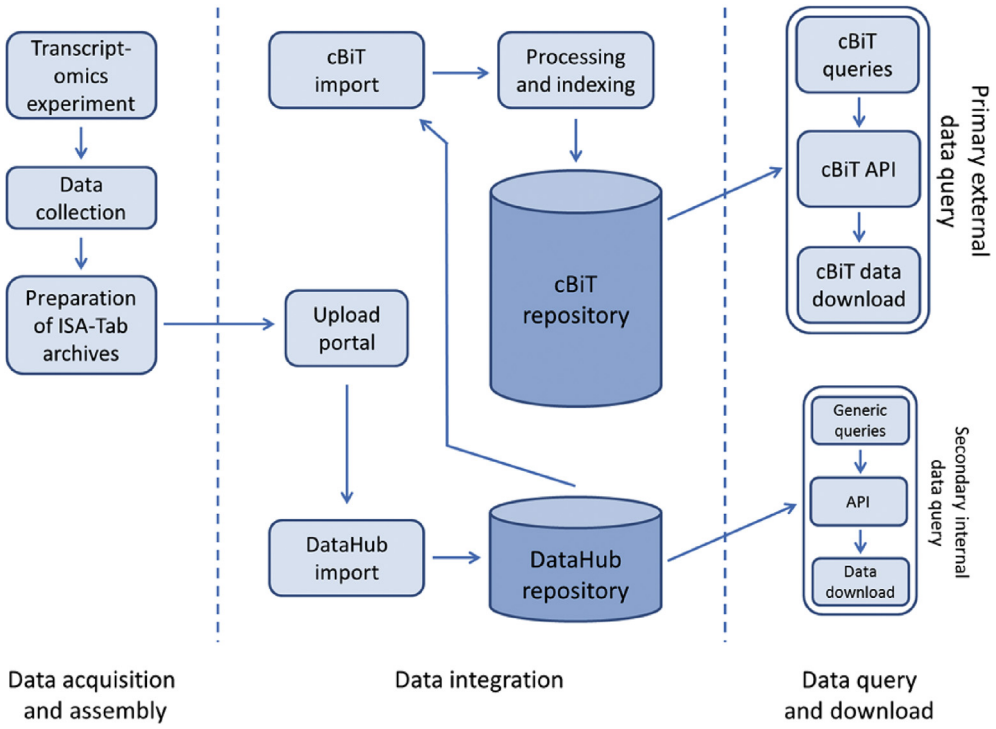

Fig. 3. Schematic overview of the cBiT data infrastructure. After data collection of a transcriptomics experiment and subsequent archive preparation, the data are first imported into the institutional DataHub repository (which also serves as a back-up) [26], followed by import into cBiT where the data are processed and indexed to enable search queries and downloads.

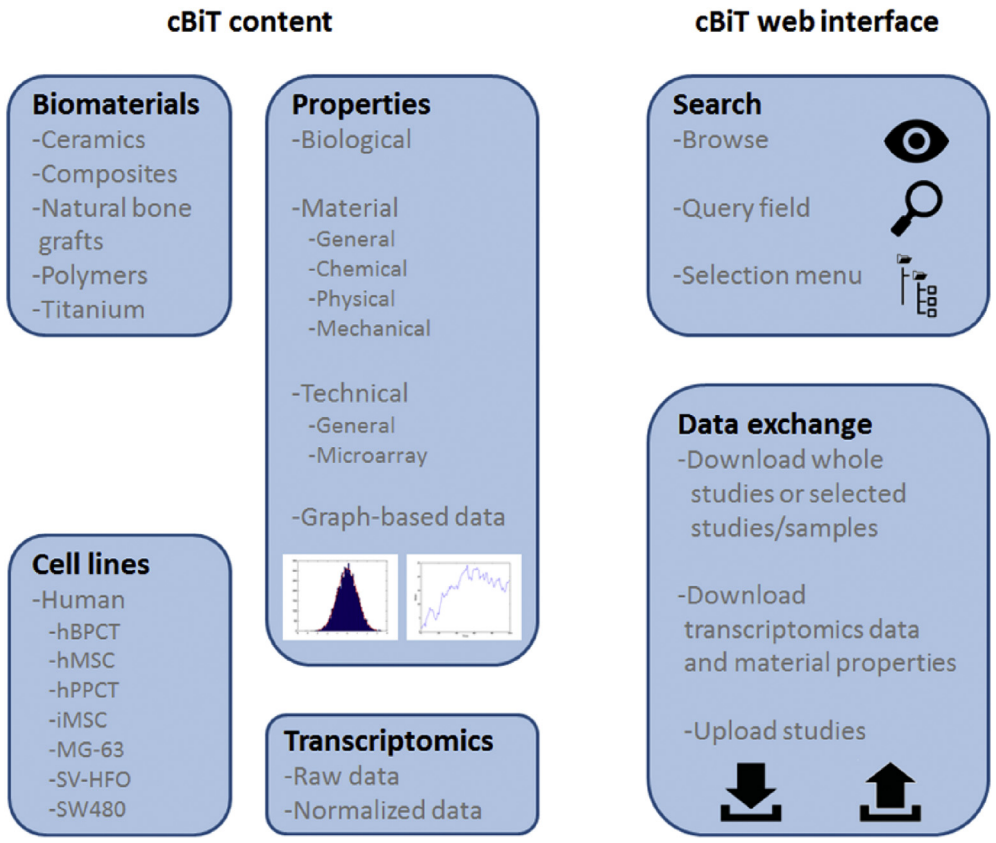

Fig. 4. An overview of cBiT's content and web interface.

transcriptomics studies. An overview of cBiT's infrastructure, from data set creation to integration and finally data download, is shown in Fig. 3.

2.1. Data sources

Currently, 8 studies including 296 samples and involving 17 different biomaterials originating from various transcriptomics projects, are deposited in cBiT. The data included in cBiT so far, have all been generated through *in vitro* microarray-based transcriptomics experiments covering 7 cell lines. An overview of what is currently covered by the data sets in cBiT is shown in Fig. 4. Each data set added to cBiT is first prepared in ISA-Tab format [27]. ISA-Tab is a tabular format, designed to be used in scientific fields to organize data into standardized data archives that distinguish between the "Investigation" (the project context), "Study" (a unit of research) and "Assay" (analytical measurement) data. This careful detailing of data archives ensures the re-usability of the data which is one of the key principles of the FAIR initiative [28]. The FAIR data principles act as an international guideline for high quality data stewardship ensuring that data are Findable, Accessible, Interoperable, and Re-usable. Other cBiT characteristics, such as its use of standard file formats, official ontology terms, and the open access structure, all contribute to meeting the FAIR principles. Furthermore, FAIR-inspired improvements, such as the addition of digital object identifiers (DOI's) to data sets, will be implemented in the near future.

cBiT was designed to include all the essential information about transcriptomics experiments as also implemented in repositories such as GEO or ArrayExpress. This involves following the MIAME (Minimum Information About a Microarray Experiment) and MINSEQE (Minimum Information about a high-throughput nucleotide SEQuencing Experiment) guidelines that enable the interpretation of the results of the experiment unambiguously and make it possible to reproduce the experiment [29]. Furthermore, we encourage the use of ISO standards such as described in ISO 10993, Biological evaluation of medical devices [30]. cBiT could thereby contribute to the goals of ISO to standardize material testing when studies incorporated in cBiT use, for example, cell types or cell lines which have been approved by ISO.

Although cBiT so far only contains microarray-based transcriptomics studies, our repository has been designed to also incorporate RNAseq data which will be included soon. Both raw and processed microarray data are available for download. RNAseq-based transcriptomics studies on the other hand will only offer the processed data since the raw data can be very large in size which could use up too much bandwidth. Instead, raw RNAseq data will be uploaded into ArrayExpress or GEO (Gene Expression Omnibus) and a link to these files will be provided within cBiT. Regarding microarray platform compatibility, cBiT can incorporate Illumina, Affymetrix, and Agilent data. For future RNAseq data, the focus will first be on the Illumina platform, but other platforms can also be included.

Data on biomaterial properties are either included in tabular format (for single numerical and categorical properties) or in a separate metadata file, containing more complex data, such as graphs or images. For instance, cBiT now contains 2 studies which incorporate graph-based data. Since samples are identically encoded in both the transcriptomics and biomaterial properties data, a download of a specific set of samples will always contain both types of data for the samples requested.

2.2. Web interface

The cBiT *Welcome* page can be found at https://cbit.maastrichtuniversity.nl and contains a short video describing the story behind cBiT. The *About* tab contains some additional information and a manual on how to use cBiT, while a Frequently Asked Questions (FAQ) list can be found in the *FAQ* tab. Exploring the content of cBiT and searching for studies with specific details can be done in the *Browse* tab. As a default, in the results section on the right side of the screen, all available studies are shown as "study cards" with the title, the authors involved, and a collapsible overview of the samples in each study. All studies can be downloaded directly, without performing any queries, using the Download button in each study card. However, if a user wants to perform a specific query, there are two ways to search through the cBiT data sets:

(1) Use the free-text search field in the upper left corner of the screen to look for specific study terms. Most study properties are indexed when new data sets are integrated into cBiT and can be queried using free-text input. The results section on the right immediately updates to display the subset of studies and samples matching the search query.
(2) Use the facetted quick selection menu on the left of the screen to narrow down studies. The menu items correspond with study properties and selecting/de-selecting items will immediately update the results section to display studies and samples that match the selection criteria. Properties are divided into Material, Biological, and Technical Properties and further sub-categories per item. An overview of all available biomaterial properties included in cBiT is shown under "Full list of properties" in the *Browse* tab.

By default, the control samples corresponding to any of the queried samples are automatically included in the search results. This function can be switched off by unchecking the "Include associated controls" check box. Next, studies or specific samples in the results section can be added to the Download box shown in the upper right corner of the screen. Clicking on this box allows the user one final check of the selection after which a zip archive is prepared which can be downloaded. The zip archive contains the raw and normalized transcriptomics data, the study protocols, the study properties, including all measured material properties, and any additional supplementary files.

2.3. Quality control and pre-processing

Data in cBiT have undergone a quality control check and a platform-specific pre-processing procedure that is also described in detail in the protocols associated with each study. Quality control is based on an evaluation of raw and normalized data density plots, cumulative distribution function (CDF) plots, MA plots, clustering and principal component analysis (PCA) plots. If these plots show any strong outliers, the data of the corresponding sample will not be included in the archive. Pre-processing of raw data consists of a platform-specific normalization and a data transformation (e.g. log2, vst, etc.).

Other research groups who are interested in uploading their study to cBiT are also invited to contact us. We have detailed instructions prepared on how to accurately prepare the data according to our standards using template files (available for download in the *About* tab). All submissions will be manually curated and any data inconsistencies will be checked and corrected in collaboration with the submitting party. Quality control of the transcriptomics data, as described above, will also be performed, thereby assuring high quality error-free data.

2.4. Data availability and permissions

Research groups who would like to integrate their data in cBiT, can already submit their data set before any accompanying manuscript is published. The data set can be put on hold and made publicly accessible once the data are published. This is in agreement with the aforementioned FAIR and MIAME/MINSEQE guidelines but also with an increasing number of journals that consider full data availability a prerequisite. The copyright of any submitted data will remain with the submitting party and any request to withdraw a data set from cBiT will be met immediately. However, it is the authors' responsibility to make sure this does not interfere with the data policy of the institution or the journal in which the manuscript was published.

With regard to data permissions, especially data based on human samples, it is the submitting party's responsibility to ensure that the submitted information does not compromise participant privacy and is in accord with the original consent in addition to all applicable laws, regulations, and institutional policies.

3. Study properties

Each cBiT study contains a wide range of material, biological, and technical properties that offer detailed measurements of specific material properties and contain biological and technical details on the biological material(s) used. We made a pre-selection of widely used properties, but these are in no way restrictive and any property can be added if required. When preparing a study for upload into cBiT, these study properties are all integrated into our standardized ISA-Tab format, ensuring comparability across studies. Where possible, official ontology terms are used to describe properties (derived from the EMBL-EBI Ontology Lookup Service, http://www.ebi.ac.uk/ols).

3.1. Material properties

Under material properties, all types of material measurements can be included. Four sub-categories exist for General, Chemical, Physical, and Mechanical properties. When building an ISA-Tab archive, the measurement device and unit of measurement are also recorded and will be part of the eventual download. Simple numerical data are included as a table, but more complex graph-based data can also be included as a supplementary Excel file.

3.2. Biological properties

The biological properties contain all relevant information on the biological material used (e.g. cell type, tissue, species, etc.) and the culturing conditions (growth medium, cell attachment, growth duration, compound exposure, etc.). Together with the study protocols included in each download, this contains all the information for a user to repeat the experiment.

3.3. Technical properties

The technical properties contain all relevant information on the transcriptomics technique used (e.g. platform, chip type) and details on the data processing such as background correction, type of normalization, and type of data transformation. Together with the study protocols included in each download, this contains all the information for a user to repeat the data processing.

3.4. Supplementary info

In addition to the optional inclusion of a supplementary Excel file for graph-based material property data, other supplementary info can also be incorporated in a study. Possibilities include RT-PCR data, Western blot images, ELISA's, colorimetric assays, etc. When available, files containing such data are part of the download.

4. Current developments

cBiT is a stable and long-term data repository that will continue to be updated with new biomaterial-based transcriptomics data archives. Our aim is to reach 30 studies in 2020. Since transcriptomics are the most actively produced omics data they are a great starting point and within a year cBiT will also be storing the first RNAseq data archives. However, in the future, cBiT can be adapted to also house other omics data, such as proteomics, epigenomics, or metabolomics data. This would require only minor modifications to the underlying data infrastructure and would make valuable new data accessible to the research community. The combined analysis of multi-omics data would also enable a true systems biology approach towards cell-biomaterial interactions.

We invite research groups who are interested in depositing their biomaterial-associated data in a sustainable repository and who are eager to contribute to the cBiT project, to contact us. It is possible for them to either deposit their own data, or provide us with biomaterials that have clinical potential which we will proceed to investigate, analyze, and incorporate into cBiT. This can also include data analysis of data already present in cBiT. In short, the possibilities offered by cBiT are:

1. Browse, download and analyze data yourself
2. Browse, download and find an external party to analyze the data
3. Contact us with an interesting question and ask us to analyze cBiT data

Table 1
List of file types typically present in a cBiT download.

File type	Description
field_descriptions.tsv	Alphabetical list of all study properties contained in the download
i_Investigation.txt	The study protocols and some additional information on associated publications
sample_info.csv	Values of all measured or documented study properties for the selected samples in the study, i.e. the metadata
Biomaterial graphs file[a]	A separate metadata Excel file, containing more complex study property data, such as graphs or images
Supplementary file[a]	One or more supplementary files containing additional measured or relevant data, not included in the sample_info.csv file or Biomaterial graphs file.
Raw data file(s)[b]	The raw data file(s) of the transcriptomics assay corresponding with the selected samples
Processed data file	The processed data file(s) of the transcriptomics assay corresponding with the selected samples and having undergone processing steps such as normalization
Annotation file	Additional annotations for the genes in the raw and processed data files (e.g. EntrezGene ID, Gene Symbol, etc.)

[a] Not always present.
[b] Not present for RNAseq data.

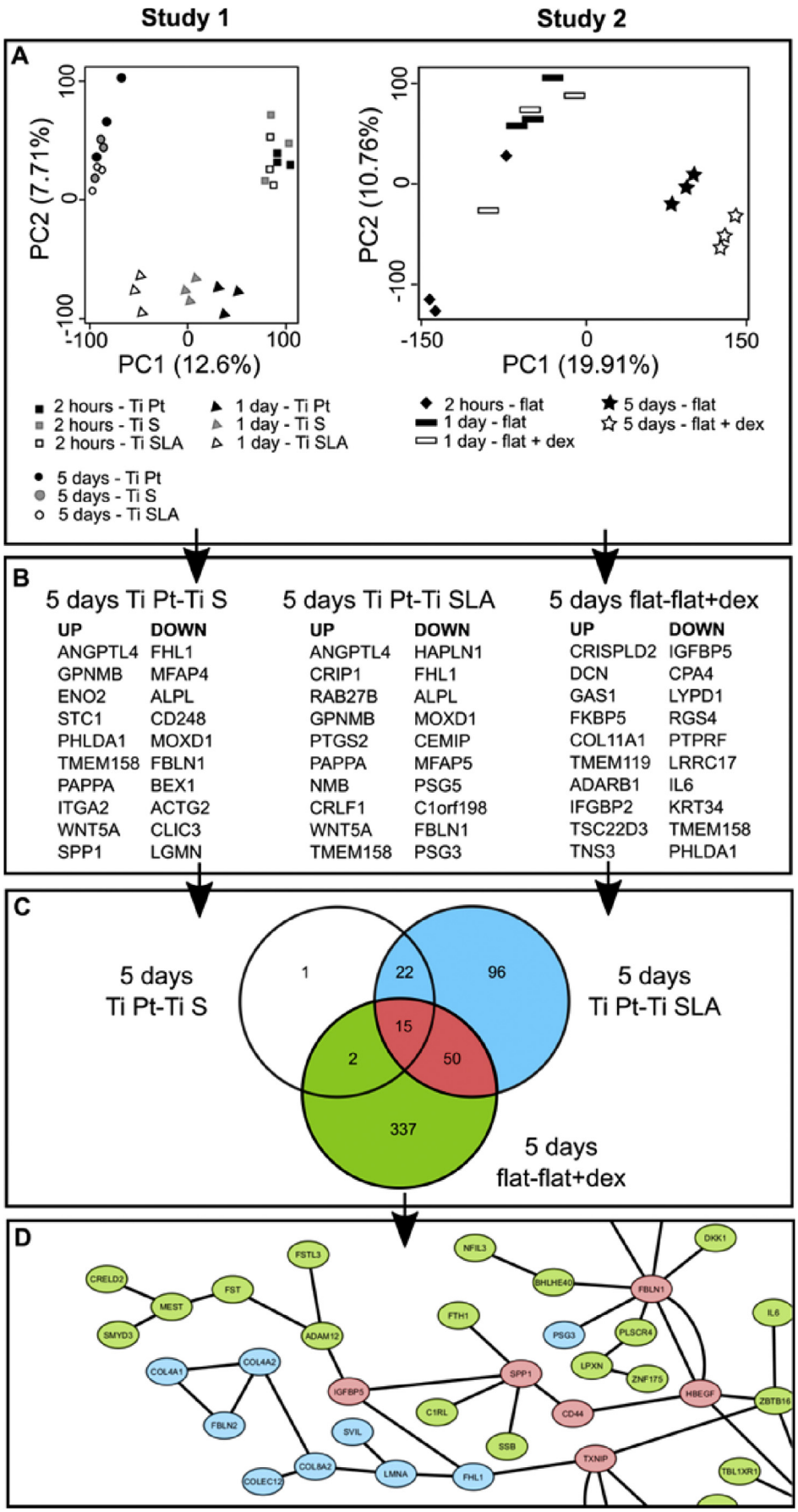

Fig. 5. A use case of cBiT data. (**A**) PCA of gene expression profiles of mesenchymal stem cells on distinct titanium surfaces. Ti = titanium (Straumann), Pt = pre-treated (flat), S = sand-blasted, SLA = sand-blasted acid etched, flat = titanium-coated flat polystyrene, dex = dexamethasone. (**B**) Top 10 of differentially expressed up- and downregulated genes on two distinct titanium surfaces. (**C**) Venn diagram of the DEGs. (**D**) Detail of interaction network based on the DEGs (full network in Supplementary Fig. S1). Blue = unique for SLA, green = unique for dexamethasone, red = overlapping between both. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Supply us with an interesting biomaterial and ask us to generate data
5. Upload your own data set

With cBiT we also aim to make a first attempt at dealing with systematic characterization of material properties. Comparison of material-based studies, even those using the same material, can be problematic due to ambiguous or inaccurate recording of material properties, used instruments or units of measurement. With our ISA-Tab archives, which require a careful documentation of all relevant details, this incertitude can be diminished. We hope that this lays the foundation for an internationally recognized and standardized biomaterial ontology.

Considering cBiT's goal of driving the discovery of useful materials or material properties that have desirable biological outcome parameters, the eventual integration of cBiT with analysis tools that assist in this identification process is obvious. For example, the Connectivity Map tool which compares expression profiles of small molecule-exposed transcriptomics data with the user's input profile could be easily adapted to screen transcriptomics data from cells grown on new potentially promising biomaterials against the biomaterial-induced expression profiles from cBiT [31].

As cBiT expands, clinical application will always be the focus point. Materials currently being used in a wide range of clinical settings are prime candidates for inclusion in cBiT. Using the knowledge accumulated in cBiT, newly developed materials can eventually find their way to the patient again.

5. Use case

To illustrate its ease of use and provide an example of cBiT's potential, we present here a use case to show how cBiT's data can be used to provide biological insight into biomaterial-cell interactions. In this scenario, we were interested in exploring the effect of titanium on osteoinductive differentiation of mesenchymal stem cells (MSCs). We started off by typing in “titanium” in the *Browse* tab's upper left search box. We then narrowed the search results down by selecting only studies using human mesenchymal stromal cells (hMSC) by selecting only that option under Biological Properties > Cell strain. We left the default option to include associated controls active. Two studies using MSCs showed up including a total of 51 samples: one study using titanium surfaces created by the dental implant manufacturer Straumann (Study 1) and one study using titanium-coated polystyrene surfaces which also includes samples exposed to dexamethasone, a positive control for osteoinduction (Study 2). Since the dexamethasone expression pattern forms an interesting osteoinduction reference profile for comparison with the Straumann gene expression patterns, we selected both studies for download by clicking on the green “+” button. At this point, the user has the option to manually remove or add specific samples in each study by clicking on the “matching samples” dropdown list in the study card. Using this possibility, in Study 2, we decided to only include the samples using a flat surface and dexamethasone and not the microstructured surface (called “Ti1018”), by unchecking the microstructure samples leaving a total of 42 samples. Next, by clicking the white/green plus button of both studies, the selected samples were added to the download box in the upper right corner and subsequently downloaded as a zip file. The unzipped file contains a separate “field_descriptions.tsv” file and two folders with the selected samples of the two studies including a small number of info files. An overview of possible file types in a typical download is shown in Table 1.

Although the downloaded data already contain processed data, users may prefer to do their own data processing. Hence, we used the raw data in this use case and processed them using the online gene expression analysis tool ArrayAnalysis (http://arrayanalysis.org/) [32]. The two studies were processed separately using the “Illumina QC and pre-processing” module. Data were already background subtracted (as indicated in the “sample_info.csv” file) and as further processing steps, quantile normalization and variance-stabilizing transformation (VST) were applied, while filtering out genes not reaching the detection threshold (i.e. detection p-value ≥ 0.01). A principal component analysis (PCA), shown in Fig. 5A, shows the separation of the different conditions.

The two processed data sets were subsequently analyzed separately using the “Statistical analysis” module in ArrayAnalysis (a linear model using the R package limma). To keep this example concise, we limited the analysis to the 5 days growth conditions in both studies. In Study 1 we selected the sandblasted (S) titanium and sandblasted-acid etched (SLA) titanium conditions and in Study 2 the flat titanium-coated surface with dexamethasone condition and compared them with their respective control conditions. As a cut-off for differential expression, an adjusted p-value (Benjamini-Hochberg) of <0.05 and an absolute fold change of >1.5 was used. The differentially expressed genes (DEGs) on the S surface, the SLA surface, and the dexamethasone-induced osteoinduction reference profile were subsequently compared.

The top 10 of differentially up- and downregulated genes for both studies and a Venn diagram with the DEGs overlap between the three conditions are shown in Fig. 5B and C respectively. The SLA surface affects almost every gene the S surface also regulates, but overall has a much stronger effect on the gene expression pattern, regulating an additional 146 genes. Interestingly, the SLA surface genes also show a considerable overlap with the genes regulated by dexamethasone (65 out of 404 genes). To visualize this overlap better we proceeded to build separate networks with the SLA and dexamethasone genes using the online tool ConsensusPathDB [33], imported the networks into CytoScape [34], merged them, and visualized the genes' origins by color as shown in Fig. 5D and Supplementary Fig. S1. A quick examination of the entire network reveals that the overlapping genes between the two networks include several multi-edge node genes which are likely to be key players in the osteogenic response. These genes include FBLN1 (Fibulin 1), HBEGF (Heparin-binding EGF-like growth factor), SPP1 (Osteopontin), PLAUS (Plasminogen activator, urokinase receptor), CXCL8 (C-X-C Motif Chemokine Ligand 8), and ASS1 (Argininosuccinate synthase 1). FBLN1, HBEGF, SPP1, PLAUS, and CXCL8 indeed have a clear relation with either osteogenesis or osteoclastogenesis [35–39]. This relationship is less clear for the ASS1 gene and therefore makes it an interesting candidate for further analysis. For the sake of brevity, we will not explore these data further here, but with this use case we have shown that a relatively simple combined analysis of cBiT data sets (time spent from data query to network construction was approximately 4 hours) can generate promising hypotheses for further testing.

6. Limitations and recommendations

cBiT offers an innovative new platform for biomaterial-based transcriptomics data and has the potential to more efficiently drive biomaterial development forward and increase our understanding of the mechanisms involved in cell-biomaterial interactions. Despite these possibilities, there are some limitations to be considered when using cBiT.

Transcriptomics data have a well-known drawback that applies to the interpretation of gene expression values. Gene expression takes place at the mRNA level and therefore does not guarantee that a working end product (i.e. a protein) will be formed that has a

measurable biological effect. Post-translational protein modifications and modulating miRNA's are two notable processes that can prevent a fully functional protein product from being formed. There are several ways to overcome these data interpretation problems though. Firstly, in follow-up assays, protein products of differentially expressed genes of interest could be detected using for example Western blotting or immunohistochemistry. However, this is only a workable solution when small numbers of proteins are being detected. A multi-omics approach that combines transcriptomics, proteomics and miRNA regulomics in a single experiment would open up the dark box between mRNA expression levels and protein end products. Although this is a very costly undertaking, it also offers possibilities for data analysis that, while very complex, are sure to result in groundbreaking insights. In addition, the added advantage of the centralized storage of transcriptomics data sets is that it facilitates meta-analyses which increase statistical power and make significant hits more likely to have a detectable biological outcome.

Other solutions can be found in the computational approaches used. For example, transcriptomics-based pathway over-representation or enrichment analyses can assist in finding affected biological processes instead of single genes. Since significantly modified pathways always contain a set of differentially expressed genes (instead of just a single gene), the likelihood of all these genes not being translated to a protein and thereby not affecting the pathway in question, becomes much less. Other computational approaches can be applied when the biological significance of the gene expression data is less important and the data are just being used as an expression profile that is characteristic of, for example, the response of cells to a certain material property. In such cases, the expression profiles can be used as biomarkers of material properties. Whether the genes expressed in these profiles are eventually translated to proteins becomes irrelevant then.

Regardless of these options though, when drawing conclusions from solely gene expression data, it is always wise to consider them as hypotheses in need of further testing and confirmation rather than biological facts.

Another limitation of cBiT concerns the (current lack of) biomaterial standardization. More specifically, in order to perform in-depth meta-analyses comparing several transcriptomics studies, it is essential that (1) crucial biomaterial parameters are characterized and (2) that the biomaterial characterization is standardized. This is presently not the case and even when these details have been recorded, comparability of similar materials between different labs is often quite poor. However, through the use of predefined fields in the cBiT user interface, we stimulate an extensive experimental characterization and a consistent description thereof. We also highly encourage cBiT users to apply available standards (developed, among others, by the American Society for Testing and Materials – ASTM) or to contribute to the standardization endeavor by developing new standards. Future developments of cBiT will focus on including standards as metadata files. Finally, meta-analyses of the data stored in cBiT can help to identify which biomaterial properties are crucial and should therefore be standardized.

7. Conclusion

We present the cBiT repository as a new tool to help researchers in finding (1) unique and standardized knowledge on the interaction of commonly used biomaterials with different cell types and (2) insight into the underlying biological responses. It is the first time that such a valuable combination of data is available for download in one central location. It is our goal to quickly expand cBiT with more data on exciting new materials, increasing its scientific value. At the MERLN Institute for Technology-inspired Regenerative Medicine in Maastricht, The Netherlands, we will continue to generate such data as part of our ongoing research but we also invite scientists from all over the world to contribute to cBiT. By becoming the go-to resource for biomaterial-associated data, we expect to make a major contribution to a more efficient development of new and better materials that show improved integration in the human body.

Competing interest statement

The authors have no competing interests to declare.

Funding

This work was supported by the Dutch province of Limburg in the LINK (FCL67723) ("Limburg INvesteert in haar Kenniseconomie") knowledge economy project and a VENI grant (number 15075) from the Dutch Science Foundation (NWO).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.biomaterials.2017.10.008.

References

[1] A.R. Bausch, U.S. Schwarz, Cellular mechanosensing: sharing the force, Nat. Mater. 12 (11) (2013) 948–949.
[2] S.W. Cranford, J. de Boer, C. van Blitterswijk, M.J. Buehler, Materiomics: an -omics approach to biomaterials research, Adv. Mater. 25 (6) (2013) 802–824.
[3] V.R. Kearns, R.J. McMurray, M.J. Dalby, 7-Biomaterial Surface Topography to Control Cellular Response: Technologies, Cell Behaviour and Biomedical Applications A2-Williams, Rachel, Surface Modification of Biomaterials, Woodhead Publishing, 2011, pp. 169–201.
[4] L. Moroni, J.H. Elisseeff, Biomaterials engineered for integration, Mater. Today 11 (5) (2008) 44–51.
[5] FDA, Breast Implant-associated Anaplastic Large Cell Lymphoma (BIA-ALCL), 2017, in: https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/BreastImplants/ucm239995.htm. (Accessed 4 July 2017).
[6] C. Greco, The Poly Implant Prothese breast prostheses scandal: embodied risk and social suffering, Soc. Sci. Med. 147 (2015) 150–157.
[7] N. Groen, J. van de Peppel, H. Yuan, J.P. van Leeuwen, C.A. van Blitterswijk, J. de Boer, Bioinformatics-based selection of a model cell type for in vitro biomaterial testing, Biomaterials 34 (22) (2013) 5552–5561.
[8] F.F.B. Hulshof, B. Papenburg, A. Vasilevich, M. Hulsman, Y. Zhao, M. Levers, N. Fekete, M. de Boer, H. Yuan, S. Singh, N. Beijer, M.A. Bray, D.J. Logan, M. Reinders, A.E. Carpenter, C. van Blitterswijk, D. Stamatialis, J. de Boer, Mining for osteogenic surface topographies: in silico design to in vivo osseointegration, Biomaterials 137 (2017) 49–60.
[9] N. Groen, H. Yuan, D.G. Hebels, G. Kocer, F. Mbuyi, V. LaPointe, R. Truckenmuller, C.A. van Blitterswijk, P. Habibovic, J. de Boer, Linking the transcriptional landscape of bone induction to biomaterial design parameters, Adv. Mater. 29 (10) (2017).
[10] N. Kasoju, H. Wang, B. Zhang, J. George, S. Gao, J.T. Triffitt, Z. Cui, H. Ye, Transcriptomics of human multipotent mesenchymal stromal cells: retrospective analysis and future prospects, Biotechnol. Adv. 35 (4) (2017) 407–418.
[11] W.M. Gallagher, I. Lynch, L.T. Allen, I. Miller, S.C. Penney, D.P. O'Connor, S. Pennington, A.K. Keenan, K.A. Dawson, Molecular basis of cell-biomaterial interaction: insights gained from transcriptomic and proteomic studies, Biomaterials 27 (35) (2006) 5871–5882.
[12] K.A. Power, K.T. Fitzgerald, W.M. Gallagher, Examination of cell-host-biomaterial interactions via high-throughput technologies: a re-appraisal, Biomaterials 31 (26) (2010) 6667–6674.
[13] P.A. Guerette, S. Hoon, Y. Seow, M. Raida, A. Masic, F.T. Wong, V.H. Ho, K.W. Kong, M.C. Demirel, A. Pena-Francesch, S. Amini, G.Z. Tay, D. Ding, A. Miserez, Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science, Nat. Biotechnol. 31 (10) (2013) 908–915.
[14] P. Komorowski, M. Walkowiak-Przybylo, B. Walkowiak, Early cell response to contact with biomaterial's surface, J. Biomed. Mater Res. B Appl. Biomater. 104 (5) (2016) 880–893.
[15] S. Neuss, B. Denecke, L. Gan, Q. Lin, M. Bovi, C. Apel, M. Woltje, A. Dhanasingh,

J. Salber, R. Knuchel, M. Zenke, Transcriptome analysis of MSC and MSC-derived osteoblasts on Resomer(R) LT706 and PCL: impact of biomaterial substrate on osteogenic differentiation, PLoS One 6 (9) (2011), e23195.
[16] W. Zhang, M.L. Wang, S. Khalili, S.W. Cranford, Materiomics for oral disease diagnostics and personal health monitoring: designer biomaterials for the next generation biomarkers, OMICS 20 (1) (2016) 12–29.
[17] D.G. Anderson, D. Putnam, E.B. Lavik, T.A. Mahmood, R. Langer, Biomaterial microarrays: rapid, microscale screening of polymer-cell interaction, Biomaterials 26 (23) (2005) 4892–4897.
[18] M. Hulsman, F. Hulshof, H. Unadkat, B.J. Papenburg, D.F. Stamatialis, R. Truckenmuller, C. van Blitterswijk, J. de Boer, M.J. Reinders, Analysis of high-throughput screening reveals the effect of surface topographies on cellular morphology, Acta Biomater. 15 (2015) 29–38.
[19] H.V. Unadkat, M. Hulsman, K. Cornelissen, B.J. Papenburg, R.K. Truckenmuller, A.E. Carpenter, M. Wessling, G.F. Post, M. Uetz, M.J. Reinders, D. Stamatialis, C.A. van Blitterswijk, J. de Boer, An algorithm-based topographical biomaterials library to instruct cell fate, Proc. Natl. Acad. Sci. U. S. A. 108 (40) (2011) 16565–16570.
[20] J. Seo, J.-Y. Shin, J. Leijten, O. Jeon, G. Camci-Unal, A.D. Dikina, K. Brinegar, A.M. Ghaemmaghami, E. Alsberg, A. Khademhosseini, High-throughput approaches for screening and analysis of cell behaviors, Biomaterials (2017).
[21] P. Cahan, H. Li, S.A. Morris, E. Lummertz da Rocha, G.Q. Daley, J.J. Collins, CellNet: network biology applied to stem cell engineering, Cell 158 (4) (2014) 903–915.
[22] B. Giotti, A. Joshi, T.C. Freeman, Meta-analysis reveals conserved cell cycle transcriptional network across multiple human cell types, BMC Genomics 18 (1) (2017) 30.
[23] A. Khan, D. Katanic, J. Thakar, Meta-analysis of cell- specific transcriptomic data using fuzzy c-means clustering discovers versatile viral responsive genes, BMC Bioinforma. 18 (1) (2017) 295.
[24] A. Schuttler, K. Reiche, R. Altenburger, W. Busch, The transcriptome of the zebrafish embryo after chemical exposure: a meta-analysis, Toxicol. Sci. 157 (2) (2017) 291–304.
[25] S.A. Sansone, P. Rocca-Serra, D. Field, E. Maguire, C. Taylor, O. Hofmann, H. Fang, S. Neumann, W. Tong, L. Amaral-Zettler, K. Begley, T. Booth, L. Bougueleret, G. Burns, B. Chapman, T. Clark, L.A. Coleman, J. Copeland, S. Das, A. de Daruvar, P. de Matos, I. Dix, S. Edmunds, C.T. Evelo, M.J. Forster, P. Gaudet, J. Gilbert, C. Goble, J.L. Griffin, D. Jacob, J. Kleinjans, L. Harland, K. Haug, H. Hermjakob, S.J. Ho Sui, A. Laederach, S. Liang, S. Marshall, A. McGrath, E. Merrill, D. Reilly, M. Roux, C.E. Shamu, C.A. Shang, C. Steinbeck, A. Trefethen, B. Williams-Jones, K. Wolstencroft, I. Xenarios, W. Hide, Toward interoperable bioscience data, Nat. Genet. 44 (2) (2012) 121–126.
[26] DataHub, DataHub, 2017. https://datahub.mumc.maastrichtuniversity.nl/. (Accessed 10 July 2017).
[27] P. Rocca-Serra, M. Brandizi, E. Maguire, N. Sklyar, C. Taylor, K. Begley, D. Field, S. Harris, W. Hide, O. Hofmann, S. Neumann, P. Sterk, W. Tong, S.A. Sansone, ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level, Bioinformatics 26 (18) (2010) 2354–2356.
[28] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.W. Boiten, L.B. da Silva Santos, P.E. Bourne, J. Bouwman, A.J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, R. Finkers, A. Gonzalez-Beltran, A.J. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.T. Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M.E. Martone, A. Mons, A.L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for scientific data management and stewardship, Sci. Data 3 (2016) 160018.
[29] FGED, The Functional Genomics Data Society (FGED), 2017. http://fged.org/projects/. (Accessed 10 July 2017).
[30] ISO, International Organization for Standardization ISO 10993: Biological Evaluation of Medical Devices, 2017. https://www.iso.org/ics/11.100.20/x/. (Accessed 8 July 2017).
[31] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J.P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S.A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, T.R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, Genes. Dis. Sci. 313 (5795) (2006) 1929–1935.
[32] L.M. Eijssen, M. Jaillard, M.E. Adriaens, S. Gaj, P.J. de Groot, M. Muller, C.T. Evelo, User-friendly solutions for microarray quality control and pre-processing on ArrayAnalysis.org, Nucleic Acids Res. 41 (Web Server issue) (2013) W71–W76.
[33] R. Herwig, C. Hardt, M. Lienhard, A. Kamburov, Analyzing and interpreting genome data at the network level with ConsensusPathDB, Nat. Protoc. 11 (10) (2016) 1889–1907.
[34] M.E. Smoot, K. Ono, J. Ruscheinski, P.L. Wang, T. Ideker, Cytoscape 2.8: new features for data integration and network visualization, Bioinformatics 27 (3) (2011) 431–432.
[35] M.A. Cooley, K. Harikrishnan, J.A. Oppel, S.F. Miler, J.L. Barth, C.J. Haycraft, S.V. Reddy, W. Scott Argraves, Fibulin-1 is required for bone formation and Bmp-2-mediated induction of Osterix, Bone 69 (2014) 30–38.
[36] C. De Fusco, A. Messina, V. Monda, E. Viggiano, F. Moscatelli, A. Valenzano, T. Esposito, C. Sergio, G. Cibelli, M. Monda, G. Messina, Osteopontin: relation between adipose tissue and bone homeostasis, Stem Cells Int. 2017 (2017) 4045238.
[37] F. Furlan, C. Galbiati, N.R. Jorgensen, J.E. Jensen, E. Mrak, A. Rubinacci, F. Talotta, P. Verde, F. Blasi, Urokinase plasminogen activator receptor affects bone homeostasis by regulating osteoblast and osteoclast function, J. Bone Min. Res. 22 (9) (2007) 1387–1396.
[38] J.L. Pathak, A.D. Bakker, P. Verschueren, W.F. Lems, F.P. Luyten, J. Klein-Nulend, N. Bravenboer, CXCL8 and CCL20 enhance osteoclastogenesis via modulation of cytokine production by human primary osteoblasts, PLoS One 10 (6) (2015), e0131041.
[39] S. Yu, Q. Geng, J. Ma, F. Sun, Y. Yu, Q. Pan, A. Hong, Heparin-binding EGF-like growth factor and miR-1192 exert opposite effect on Runx2-induced osteogenic differentiation, Cell Death Dis. 4 (2013), e868.