Proteomics-based Dissection of Human Endoderm Progenitors by Differential Cell Capture on Antibody Array*§

Revital Sharivkin‡, Michael D. Walker‡, and Yoav Soen‡§

Heterogeneity, shortage of material, and lack of progenitor-specific cell surface markers are major obstacles to elucidating the mechanisms underlying developmental processes. Here we report a proteomics platform that alleviates these difficulties and demonstrate its effectiveness in fractionating heterogeneous cultures of early endoderm derived from human embryonic stem cells. The approach, designated differential cell-capture antibody array, is based on highly parallel, comparative screening of live cell populations using hundreds of antibodies directed against cell-surface antigens. We used this platform to fractionate the hitherto unresolved early endoderm compartment of CXCR4+ cells and identify several endoderm (CD61+ and CD63+) and non-endoderm (CD271+, CD49F+, CD44+ and B2M+) sub-populations. We provide evidence that one of these sub-populations, CD61+, is directly derived from CXCR4+ cells, displays characteristic kinetics of emergence, and exhibits a distinct gene expression profile. The results demonstrate the potential of the cell-capture antibody array as a powerful proteomics tool for detailed dissection of heterogeneous cellular systems. Molecular & Cellular Proteomics 11: 10.1074/mcp.M111.016840, 586–595, 2012.

Although much has been learned over the past 20 years about organogenesis in endoderm-derived tissues, earlier stages of endoderm development remain incompletely understood (1, 2). Studies of gene expression and cell division rate within anterior and posterior endoderm suggest that regional identity is established already at gastrulation (3–7). Dissecting these events in a human model requires isolation of lineage-specific precursors underlying the multi-step progression of early endoderm development. Although human embryonic stem cells (hESC)1-based models of endoderm differentiation may provide a powerful model for these studies (8–11), relevant analysis is often confounded by tissue heterogeneity and insufficient numbers of precursors for screening by flow cytometry. In addition, very few markers, particularly cell-surface markers, are currently associated with specific subsets of early stage precursors in the endoderm lineage. Consequently, studies involving differentiation of hESCs toward endoderm, often categorize stage-specific cells based on the stages of differentiation protocols, overlooking the multiple cell identities that populate these cultures.

Recent studies in hESC-derived endoderm cultures have nonetheless begun to uncover cell surface markers for isolation of pancreatic endoderm-stage (12) or primitive gut tube-stage cells (13, 14). Characterization of precursor composition in the preceding stage of differentiation toward endoderm is, however, still lagging. Endoderm cells at this stage are typically identified by the expression of CXCR4, which has been correlated in mouse ES-derived cultures with definitive endoderm (15). Indeed, CXCR4 was shown to be expressed in hESC-derived cells that have been induced to differentiate toward early endoderm (16). Still, the extent of heterogeneity within CXCR4+/- compartments and the timing of emergence of additional sub-populations are unknown. Recent work in chick embryos showed that early stage CXCR4+ cells contain, in addition to endoderm cells, a small population of non-endoderm cells which contribute to the development of endoderm tissues, specifically the pancreas (17). Such studies emphasize the fundamental importance of resolving the different subsets of CXCR4+ cells of the early, definitive endoderm stage.

We wished to exploit the potential of antibody arrays to identify subsets of endoderm and non-endoderm cells appearing during early definitive endoderm development. Antibody arrays are typically used to measure the levels of proteins in cell lysates in a wide range of experimental systems (18–20; reviewed in 21). They are also used extensively in diagnostic applications, e.g. detection of biomarkers in serum (22, 23) or urine samples (24). To a lesser extent, antibody arrays have been applied to profiling cell surface markers in several normal and disease settings, such as rat neural stem cells (25) and different infectious and neoplastic disease states. These include HIV (26), leukemias (27), and colorectal cancer leukemia (28). Because these assays are based on binding of a single population to a single array, their ability to evaluate differences between populations may be limited.

1 The abbreviations used are: hESCs, human embryonic stem cells; DE, definitive endoderm; RMA, robust multi-array average.
Here we describe a novel antibody array platform termed differential cell-capture antibody array: this approach permits direct comparison of cell surface marker profiles in different populations, thereby allowing efficient identification of differentially expressed markers. The ability to compare two populations on a single array is crucial for discriminating relatively similar populations exhibiting expression changes that are subtle, rather than all-or-none. This is of particular importance for embryonic stem cell-based research where there is a need to resolve emerging precursors that may initially be quite similar. Indeed, using this approach, we have been able to efficiently identify cell surface markers expressed selectively on endoderm and non-endoderm populations of differentiating hESCs. Furthermore, use of these markers now permits sub-fractionation of the early endoderm compartment.

Molecular & Cellular Proteomics 11.9 587

**Array Fabrication**—Arrays were printed in a Microgrid printer with solid dots (Total array Systems, BioRobotics, Cambridge, UK) on hydrogel coated slides (Full Moon Biosystems, Sunnyvale, CA) using a panel of 235 monoclonal mouse anti-human antibodies (BD biosciences). The antibodies were printed at a concentration of 0.5 mg/ml in five spots, each using a single stamp and with 750 μm spacing. Following printing, the arrays were hydrated in a humidifier at 4 °C for 48 h, and then dried for 10 min at room temperature.

**Cell Preparation**—Flow cytometry isolated cells were incubated with either DIO or DID dyes (Biotium) for 45 min at 37 °C using 5 μl of dye/1 ml of hESC growth medium for every 10^6 cells. Labeled cells were washed in PBS and centrifuged at 400 x g for 5 min. Differentially labeled populations were mixed at a ratio of 1:1 at a total concentration of ~5*10^6 cells/ml in 250–500 μl of hESC growth medium, supplemented with 1 μl of DNase (Ambion 2U/μl). Prior to incubation of cells on the array, the printed area was blocked for 3 min with 1% bovine serum albumin in PBS solution. The blocking solution was replaced by the cell suspension, and the arrays were incubated for 1 h at 37 °C. Excess cells were removed in a large volume of PBS and the arrays were fixed in 4% paraformaldehyde solution for 10 min. Arrays were imaged using automated, high-content fluorescence microscopy (IXmicro, MDC).

**Real-time Quantitative PCR**—RNA from sorted populations of cells was isolated using RNeasy MiniElute Cleanup kit (Qiagen 74204). DNA was eliminated using TURBO DNA-free kit (Ambion AM1907) and the mRNA was converted to cDNA using high-capacity cDNA Reverse Transcription kit (Ambion 4374967). Transcript levels were measured using real-time qPCR on a 7900HT Fast Real-Time PCR System using Power SYBR green PCR master mix (Applied Biosystems). Primer sequences are detailed in supplemental Table S4. The levels of each gene was normalized using RPLPO as an endogenous control mRNA.

**DNA Microarray Analysis**—GeneChip (Affymetrix) human genome ST1 series arrays were used to identify genes differentially expressed between CD61+ cells and CXCR4+ cells. RNA sample preparation, array hybridization, array washing and scanning were performed according to the manufacturer’s protocol. Data from the individual arrays were acquired using GCOS software and expression values were normalized by Expression Console software using robust multi-array average (RMA) algorithm (31). Heat map analysis was performed using the “Cluster” and “TreeView” programs (http://www.eisenlab.org/eisen/?page_id=42) following subtraction of (log-scale) values measured in undifferentiated hESCs.

**Statistics**—Statistical analysis of differential capture by specific antibodies (Table I) was performed as follows: In each array experiment, we measured the percentages of CXCR4+ and CXCR4− cells bound to a specific choice of antibody spots. Repeating this scoring for independent replicate independent experiments yielded sets of percentages corresponding to a given population (CXCR4+ or CXCR4−), and a specific cell surface marker. We then used paired t-test to evaluate the likelihood that the percentages of different cell populations are drawn from the same distributions (null hypothesis). This procedure does not take into account the probability of measuring differences between the two cell populations in a given replicate (a different null hypothesis). Thus, the paired t-test underestimates the actual significance.

p values of gene expression differences (Figs. 2B; 2C; Fig 3; Fig 4B) were computed using the paired t test with two-samples for means with different variances (one-tail). p values for the significance of enrichment of vasculogenesis, smooth muscle, and ECM genes in CD61+ cells as measured by DNA micro-arrays were obtained using the DAVID web tool (DAVID bioinformatics resources 6.7) (32, 33).
RESULTS

Development of Antibody Arrays for Differential Capture of hES-derived Cells—Developing a cell-capture antibody array procedure for identification of progenitors in cultures of differentiating hESCs presents significant challenges. In particular, culture conditions during differentiation may cause significant cell stress and death. These cultures tend to be heterogeneous and may include sub-populations that are similar and difficult to resolve. Thus, the antibody array platform must be sensitive enough to detect small subpopulations of progenitors and specific enough to avoid background binding of a variety of cell types that may exist in these cultures. Moreover, practical use of the array with differentiating hESCs requires procedures for reducing cellular stress which could otherwise compromise cell capture on the array. We achieved these goals by substantially modifying an approach that we have previously employed for identifying antigen specific T-cells using printed arrays of peptide-MHC (34, 35). Of particular importance were the choice of slide coating, the choice of antibodies and their deposition on the array, and the cellular handling prior to analysis on the array (elaborated in the Supplementary information). Combining these modifications with direct comparison of the two populations on a single array was essential for discriminating relatively similar populations. The modified platform and procedures made it possible to reliably identify endoderm and non-endoderm markers in differentiating populations of hESCs and resolve sub-populations within these compartments.

We generated arrays containing 235 different antibodies (all flow-cytometry validated; supplemental Table S1), printed with 5 replicates each, on film-coated glass slides. We used these arrays for differential capture of two distinct cell populations, each labeled with a different nonspecific membrane dye (DIO and DID). Labeled cells were mixed in equal numbers and applied in solution to the antibody array. Capture on specific antibody spots was based on recognition of cell-surface antigens by the printed antibodies (Fig. 1A). Following removal of unbound cells, the array was imaged using automated, high content fluorescence microscopy (ImageXpress Micro) and scored for the number of cells of each color on each spot. Spots that were enriched for cells of one color identified candidate cell surface markers characteristic of the respective cell population. Since the entire capture assay is performed on a small surface of a single film-coated slide, it is readily applicable for samples with limited number of cells, for which flow cytometry screens are impractical.

Identification of Surface Markers Distinguishing CXCR4+ from CXCR4– Cells—We used the array to screen for differentially expressed cell-surface markers characteristic of hESC-derived endoderm and non-endoderm cells. The initial distinction between these populations was based on the expression of CXCR4 (15). hESCs (HUES-2 line (29)) were exposed to a 3-day treatment promoting differentiation toward definitive endoderm (DE) (30). We isolated CXCR4+ and CXCR4– cells by flow-cytometry, labeled 2–5*10^5 cells of each population respectively, with DIO (green) and DID (far red) membrane dyes, and analyzed differential binding of the labeled populations to antibody spots (Fig. 1B; Table I). As expected, spots printed with anti-CXCR4 (α-CD184) repro-
Dissection of Human Endoderm Progenitors by Antibody Array

**Table I**

Cell-surface markers detected on endoderm and non-endoderm cells. A. Summary of results from four independent antibody array analyses using CXCR4+/− populations derived from the HUES-2 cell-line. Shown are markers detected in two or more experiments. B. Differentially expressed markers identified by the array (including CXCR4). Top and bottom sections display markers preferentially expressed by the CXCR4+ and CXCR4− populations, respectively. The criteria used for determining differential binding were: (1) a greater than 20% difference in percentage of bound cells and, (2) p value < 0.02. p values are based on paired t-test as described in the experimental procedures section. Percentages of binding in each replicate are displayed in supplemental Table S2.

**A. Detected markers**

| Marker     | Captured cells (sum over 5 spots) | CXCR4+ : CXCR4− (%) | p value (<) |
|------------|-----------------------------------|---------------------|-------------|
| CXCR4      | 684                               | 92 : 8              | 7.5e-6      |
| CD63       | 595                               | 63 : 37             | 0.014       |
| CD61       | 81                                | 87 : 13             | 0.0016      |
| CDW93      | 422                               | 73 : 27             | 0.012       |
| CD15       | 174                               | 73 : 27             | 0.019       |
| CD140B     | 162                               | 86 : 14             | 6.5e-4      |
| CD49F      | 362                               | 29 : 71             | 0.012       |
| CD44       | 167                               | 26 : 74             | 0.01        |
| CD271      | 413                               | 25 : 75             | 0.04        |
| B2M        | 103                               | 31 : 69             | 0.012       |

**B. Differentially expressed markers**

| Marker     | Captured cells (sum over 5 spots) | CXCR4+ : CXCR4− (%) | p value (<) |
|------------|-----------------------------------|---------------------|-------------|
| CXCR4      | 684                               | 92 : 8              | 7.5e-6      |
| CD63       | 595                               | 63 : 37             | 0.014       |
| CD61       | 81                                | 87 : 13             | 0.0016      |
| CDW93      | 422                               | 73 : 27             | 0.012       |
| CD15       | 174                               | 73 : 27             | 0.019       |
| CD140B     | 162                               | 86 : 14             | 6.5e-4      |
| CD49F      | 362                               | 29 : 71             | 0.012       |
| CD44       | 167                               | 26 : 74             | 0.01        |
| CD271      | 413                               | 25 : 75             | 0.04        |
| B2M        | 103                               | 31 : 69             | 0.012       |

Ducibly captured CXCR4+ (green-labeled) cells, but not CXCR4− (red-labeled) cells (Fig. 1B top row, center; Table 1B top row). Additionally, the analysis identified 51 markers (Table 1A) expressed by the populations of interest, ~75% of which were detected in all 4 replicates of the experiment. Analysis of differential capture identified several antibodies which preferentially captured CXCR4+ cells (e.g. CD61 and CD63; Fig. 1B, top row) and antibodies which preferentially recognized CXCR4- cells (e.g. CD49F and CD271; Fig. 1B, middle row). The reproducible identification of expressed and differentially expressed markers demonstrates the efficiency of this platform parallel and rapid screening of hundreds of surface markers using a limited number of cells (typically ~10^6 cells per profiling assay).

Validation of Differentially Expressed Markers—We validated the outcome of the cell capture assay by two independent methods: flow cytometry and RT-qPCR. To verify the association of the new surface markers with CXCR4, we induced differentiation of hESCs toward DE, co-stained the cells with CXCR4 and each of the identified markers, and analyzed the staining by flow cytometry (Fig. 2A). Most of the differential surface markers that were detected by the array exhibited the expected correlation with the expression of CXCR4. For example, higher levels of CD61 and CD63 were detected in CXCR4+ cells, whereas CD49F and CD271 showed higher expression in CXCR4− cells (Fig. 2A). Similar results were obtained using the hESC line H9 (supplemental Fig. S1). Importantly, cells expressing high levels of the newly identified markers comprised only a subset of the CXCR4+ (or CXCR4−) population, suggesting that these markers designate distinct sub-populations of each compartment.

Using RT-qPCR, we found that CXCR4+ cells indeed express significantly higher mRNA levels of markers that were preferentially detected by the antibody array in CXCR4+ cells, and vice versa for CXCR4− cells. Specifically, CD61 and CD63 mRNAs were elevated in the CXCR4+ cells, whereas CD49F, CD271, CD44, and B2M expression levels were higher in CXCR4− cells (Fig. 2B). In a reciprocal test, we isolated cells by flow cytometry using several of the identified surface markers and analyzed the levels of CXCR4 mRNA. These levels were in accordance with the differential expression of the identified markers in the CXCR4+ or CXCR4− population. For example, CXCR4 expression was higher in CD61+ cells compared with CD61− cells, but lower in CD271+ and CD49F+ cells versus their inverse, CD271− and CD49F− populations (Fig. 2C). Thus, flow cytometry- and mRNA-based analysis of these markers confirmed the array-based results and demonstrated the fidelity of the array in discovering bona fide selectively expressed markers.

CD61+ Cells Exhibit an Endoderm Signature with Distinct Gene Expression Characteristics—Next, we tested whether the newly identified populations display the expected endoderm or non-endoderm characteristics. RT-qPCR analysis of expression revealed that CD61+ cells express significantly higher levels of the endoderm genes, FOXA2, SOX17, and
HNF1β compared with CD61– cells (Fig. 3). Conversely, CD49F+ and CD271+ cells expressed much lower levels of the endoderm genes and higher levels of non-endoderm genes (e.g. Brachyury) compared with their inverse fractions (Fig. 3 and supplemental Fig. S2, respectively).

Although the CD61+ population shared the endoderm characteristics of CXCR4+ cells, it comprised only a small subset of the CXCR4+ compartment (Fig. 2A, top left). This suggests that CD61+ cells might also have characteristics which distinguish them from the remaining CXCR4+ cells. To test this, we compared the genome-wide mRNA profile of CD61+ cells to that of CXCR4+ cells using Affymetrix arrays.

Overall, the profiles of CXCR4+ and CD61+ cells were very similar (Fig. 4A, left), especially with respect to endoderm genes; indeed, none of the classical endoderm genes were differentially expressed in CD61+ compared with CXCR4+ cells. Despite this similarity, CD61+ cells exhibited a distinct signature of ~100 genes that were expressed at higher levels compared with CXCR4+ cells (Fig. 4A, right; supplemental Table S3). Gene ontology analysis (using two different functional annotation programs) revealed that this signature includes significant enrichment of genes involved in vasculogenesis/angiogenesis (p value 3E-7), smooth muscle genes (p value 1E-7), and ECM genes (p value 7E-13). We validated the
e.g. EPAS1, CAV1, TNC, Serpine1 genes,[24] representative endothelial and muscular sub-populations. Real-time qPCR analysis of differential expression of endoderm and non-endoderm genes in FACS-sorted populations of CD61+/H11001 and CD49F+ versus CD49F− (n = 4). Values shown are mean ± S.E. (* denotes p value <0.05, ** denotes p value <0.01).

differential expression of representative endothelial and muscle genes (e.g. EPAS1, CAV1, TNC, Serpine1, and ACTA2) by RT-qPCR and indeed, found elevated expression in isolated CD61+ cells compared with CXCR4+ (Fig. 4B) and CXCR4+/CD61− cells (supplemental Fig. S3). Thus CD61+ cells represent a sub-population of CXCR4+ cells with a distinct gene expression profile.


**CD61+ Cells Show Specific Kinetics of Appearance and Are Generated From CXCR4+ but not CXCR4− Cells**—To compare the kinetics of appearance of CD61+ cells and CXCR4+ cells, we analyzed the temporal expression pattern of both populations during 6 days of endoderm differentiation in response to activin A. Over this time period, we measured the fraction of positive cells by flow cytometry and the mRNA levels of relevant genes by RT-qPCR. We found a small fraction of CXCR4+ cells already at day 1, and a dramatic increase in CXCR4+ cells that typically occurred on day 2 or 3 (Fig. 5A). CD61+ cells emerged a day later and their fraction typically increased on day 3 or 4 (Fig. 5A). Likewise, the increase in CXCR4 mRNA (as well as the endoderm marker SOX17) clearly preceded that of CD61 and the endothelial gene EPAS1 (Fig. 5B–5C). Similar results were obtained in 3 independent flow cytometry and RT-qPCR experiments, and are consistent with the idea that CD61+ cells are derived from CXCR4+ cells. To test this, we isolated CXCR4+/CD61− cells at day 3 of differentiation and examined the expression of CD61 following two more days of spontaneous differentiation. Analysis by flow cytometry revealed that CD61+ cells reproducibly emerged from CXCR4+/CD61− cells (Fig. 6A), but not from the CXCR4− cells (supplemental Fig. S4). As expected, the emergence of CD61+ cells in differentiating cultures of CXCR4+/CD61− cells was accompanied by up-regulation of CD61 mRNA (Fig. 6B; supplemental Fig. S5). Although we cannot exclude the possibility that some CD61+ cells are present in the CXCR4+/CD61− cell population, such residual cells are unlikely to explain the observed accumulation of CD61+ cells over time. Indeed, the kinetics of emergence of CD61+ cells (Fig. 5A) did not indicate that these cells grow rapidly, suggesting that the re-appearance of CD61+ cells in a CXCR4+/CD61− culture is probably not the result of differential expansion. These results indicate that the CD61+ population is derived from the preceding stage of CXCR4+ cells.

**DISCUSSION**

In this study, we introduce an efficient, proteomics-based method for resolving cellular heterogeneity in samples of stem and progenitor cells derived from hESCs. We used this platform to: (1) identify several surface markers discriminating early hESC-derived endoderm from non-endoderm cells, and (2) to subfractionate the definitive endoderm population of CXCR4+ cells into discrete subsets exhibiting unique properties. The analysis revealed a hitherto unrecognized heterogeneity within the early endoderm compartment, and identified sub-populations of endoderm (CD61+, CD63+) and non-endoderm cells (CD271+, CD49F+, CD44+, B2M+) that emerge in response to endoderm-promoting treatment.

The CD61+ subset represents one such sub-population that was generated from CXCR4+ cells. It exhibited clear endoderm features but included, in addition, a characteristic signature of genes which distinguished it from the bulk of CXCR4+ cells. The CD61+ population also exhibited distinct kinetics of appearance; its emergence and peak expression lagged behind the CXCR4+ population by 1 day. Throughout this time, CD61+ cells comprised less than 20% of the CXCR4+ population, indicating that CD61+ cells co-exist with the remaining CXCR4+ cells and do not simply correspond to the next stage of definitive endoderm development. The emergence of a distinct population of CD61+ cells from CXCR4+/CD61− endoderm cells provides initial delineation of early stages in the developmental scheme of the definitive endoderm lineage. Interactions between these populations of CXCR4+/CD61+ and CXCR4+/CD61− cells may play a significant role in further patterning of the CXCR4+ compartment.

Resolving distinct sub-populations within heterogeneous cultures of cells is a major technical challenge, especially when the number of cells is limited and the populations are similar. Previous platforms and procedures involving antibody arrays were not tailored for such analysis. Here we achieved this goal by developing a differential cell-capture assay that

**Fig. 3.** Gene expression characteristics of endoderm (CD61+) and non-endoderm (CD49F+) sub-populations. Real-time qPCR analysis of differential expression of endoderm and non-endoderm genes in FACS-sorted populations of CD61+ versus CD61− and CD49F+ versus CD49F− (n = 4). Values shown are mean ± S.E. (* denotes p value <0.05, ** denotes p value <0.01).
enables direct comparison of cell-surface antigen profiles of different cell populations, thus permitting discrimination between different sub-populations of cells. Starting with a single distinguishing antigen, CXCR4, we were able to identify a set of additional markers that are associated with either CXCR4+/H11001 or CXCR4− cells. Each of these additional markers is itself a potential identifier of a sub-population thus, enabling sub-fractionation of the CXCR4+/H11001 and CXCR4− compartments. In addition, these markers can be used in combinations to further purify the progenitors at higher resolution.

**Fig. 4.** CD61+ cells exhibit a distinct gene expression profile. A, Clustered heat map representation of gene expression fold-change in CD61+ and CXCR4+ cells versus undifferentiated hESCs as measured by DNA micro-arrays. Overall, CD61+ and CXCR4+ cells exhibited a very similar global pattern of gene expression (left). However, the profile of CD61+ cells included a group of genes that were elevated compared with CXCR4+ cells (right), some of which are indicated in the expanded view. Color bar scales (log2) are displayed at the bottom of each panel. The comprehensive gene expression signature of CD61+ cells is provided in supplemental Table S3. B) Real-time qPCR validation of several genes distinguishing CD61+ from CXCR4+ endoderm cells. Values shown are mean ± S.E. (n = 7, * denotes p value <0.05, ** denotes p value <0.01).
Our cell-capture antibody array approach proved to be sensitive, accurate and reproducible in a fluctuating cellular context. The sensitivity is manifested by the ability to detect small populations such as CD61/H11001 (constituting ~2–5% of the population). The exact detection limit however, may vary according to the abundance of the antigen on the surface of positive cells, the properties of the respective antibody, the total number of cells deposited on the array, and the degree of nonspecific binding to the array. The nonspecific binding is greatly reduced by the hydrogel slide coating, thus contributing to high signal-to-noise ratio compared with other substrates. Owing to the low background binding, the sensitivity for a particular case of marker and antibody can be improved by increasing the number of replicate spots or the number of cells deposited onto the array. In our experiments, we analyzed as few as 0.4 × 10⁶ cells with five replicated spots per antibody. Reducing the number of cells would limit the ability to detect rare populations of positive cells and would therefore compromise sensitivity.

The reproducibility of the differential cell-capture approach is affected by both the performance of the array and the variability of marker expression in the sample. In the case of hES-derived cells, the intrinsic variability of the differentiation outcome is often quite high. Nonetheless, 75% of the markers were detected in all four replicate experiments, indicating high degree of reproducibility. Judging from estimations of accuracy (below), we suspect that most of the variability is due to biological variation, as opposed to variation introduced by the array procedure.

The accuracy of the array was evaluated by the false positive rate of detection of differential markers, as judged by validation using FACS (Fig. 2A) and qPCR (Fig. 2B). In our hands, only one of 10 differential markers failed to pass the validation in the follow-up experiments, indicating a false positive rate of about 10%.

Overall, the comparative array approach permits direct discrimination of surface markers characteristic of different cell populations, thus providing a powerful tool for enhancing the identification of developmental stage-specific progenitors. Although we demonstrated the usefulness of this tool for hESC-based research, it is as applicable for dissecting in-vivo samples. It therefore constitutes a general and efficient platform for resolving complex developmental programs.

**Fig. 5. CD61+ cells exhibit distinct temporal characteristics.** A, Flow cytometry analysis of the kinetics of appearance of CXCR4+ cells (green) and CD61+ cells (blue) during 6 days of differentiation with Activin A. Flow cytometry plot color code is as in Fig. 2A. Percentage of each population (relative to the total number of live cells) is indicated within each panel (green and blue squares to indicate CXCR4+ and CD61+ cells, respectively). Shown are live cells as determined by cell size and PI staining. B, Complementary kinetics analysis of CXCR4 versus CD61 mRNA expression measured by real-time qPCR. C, Similar real-time qPCR analysis of mRNA expression of SOX17 and EPAS1.
Dissection of Human Endoderm Progenitors by Antibody Array

**A**  
Flow cytometry analysis of isolated, hESC-derived CXCR4+/CD61- cells and CXCR4+/CD61+ cells (dark green) following 3 days of differentiation with activin A (top panel). The bottom panel demonstrates re-emergence of CD61+ cells in the isolated CXCR4+/CD61− fraction following two additional days of spontaneous differentiation in culture. Expression of CD61 is indicated by APC labeling. Shown are live cells as determined by cell size and PI staining. **B**  
Real-time qPCR analysis of the experiment shown in (A). CD61 and CXCR4 expression in isolated CXCR4+/CD61− cells after 2 days in culture relative to their expression prior to the incubation. At the time of isolation, the absolute level of CD61 mRNA in CXCR4+/CD61− fraction was ~130-fold lower than that of CXCR4 in these cells.

**Fig. 6.** CD61+ endoderm cells are derived from CXCR4+ cells. A. Flow cytometry analysis of isolated, hESC-derived CXCR4+ (light green) and CXCR4+/CD61− cells (dark green) following 3 days of differentiation with activin A (top panel). The bottom panel demonstrates re-emergence of CD61+ cells in the isolated CXCR4+/CD61− fraction following two additional days of spontaneous differentiation in culture. Expression of CD61 is indicated by APC labeling. Shown are live cells as determined by cell size and PI staining. **B.** Real-time qPCR analysis of the experiment shown in (A). CD61 and CXCR4 expression in isolated CXCR4+/CD61− cells after 2 days in culture relative to their expression prior to the incubation. At the time of isolation, the absolute level of CD61 mRNA in CXCR4+/CD61− fraction was ~130-fold lower than that of CXCR4 in these cells.

Acknowledgments—We thank Michael Elgart for valuable assistance in generating analysis scripts, Prof. Nissim Benvenisty and Prof. Michael Fainzilber for helpful advice, and Dr. Shirley Horn-Saban from the Weizmann Microarray Facility for technical support. We also thank the members of the Soen and Walker groups for their constructive inputs.

* This study was supported by a network grant (to Y.S. and M. W.) from the Juvenile Diabetes Research Foundation (JDRF) and by the Leona M. and Harry B. Helmsley Charitable Trust (to Y.S.). Y.S. was supported by a CDA award from the Human Frontier Science Program (HFSPO). Y.S. is Incumbent of the Daniel E. Koshland Sr. Career Development Chair at the Weizmann Institute. M.W. is Incumbent of the Marvin Meyer and Jenny Cyker Chair of Diabetes Research at the Weizmann Institute.

This article contains supplemental Figs. S1 to S5 and Tables S1 to S4.

1. *To whom correspondence should be addressed: Dept of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. Tel: +972-8-934-6011; Fax: +972-8-934-4118; E-mail: yoavs@weizmann.ac.il.*

REFERENCES

1. Grapin-Botton, A., and Melton, D. A. (2000) Endoderm development: from patterning to organogenesis. Trends Genet. **16**, 124–130

2. Zorn, A. M., and Wells, J. M. (2009) Vertebrate endoderm development and organ formation. *Annu. Rev. Cell Dev. Biol.* **25**, 21–251

3. Katsumoto, K., Fukuda, K., Kimura, W., Shimamura, K., Yasugi, S., and Kume, S. (2009) Origin of pancreatic precursors in the chick embryo and the mechanism of endoderm regionalization. *Mech. Dev.* **126**, 539–551

4. Lawson, K. A., and Pedersen, R. A. (1987) Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* **101**, 627–632

5. Thomas, P. Q., Brown, A., and Beddington, R. S. (1998) Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development* **125**, 85–94

6. Tremblay, K. D., and Zaret, K. S. (2005) Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. *Dev. Biol.* **280**, 87–99

7. Burtscher, I., and Lickert, H. (2009) Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029–1038

8. Ameri, J., Stahlberg, A., Pedersen, J., Johansson, J. K., Johannesson, M. M., Artner, I., and Semb, H. (2010) Foxg2 specifies hESC-derived definitive endoderm into foregut/midgut cell lineages in a concentration-dependent manner. *Stem Cells** **28**, 45–56

9. Kopper, O., Giladi, O., Golan-Lev, T., and Benvenisty, N. (2010) Characterization of gastrulation-stage progenitor cells and their inhibitory crosstalk in human embryoid bodies. *Stem Cells** **28**, 75–83

10. Séguin, C. A., Draper, J. S., Nagy, A., and Rossant, J. (2008) Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* **3**, 182–195

11. Yu, P., Pan, G., Yu, J., and Thomson, J. A. (2011) Foxg2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation. *Cell Stem Cell* **8**, 326–334

12. Jiang, W., Sui, X., Zhang, D., Liu, M., Ding, M., Shi, Y., and Deng, H. (2011) CD24: a novel surface marker for PDX1-positive pancreatic progenitors derived from human embryonic stem cells. *Stem Cells** **29**, 609–617

13. Lee, D. H., Ko, J. J., Ji, Y. G., Chung, H. M., and Hwang, T. (2011) Proteomic Identification of RREB1, PDE6B, and CD209 Up-Regulated in Primitive Gut Tube Differentiated From Human Embryonic Stem Cells. *Pancreas** **41**, 65–73

14. Wang, P., Rodriguez, R. T., Wang, J., Ghodasara, A., and Kim, S. K. (2011) Targeting SOX17 in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. *Cell Stem Cell** **8**, 335–346

15. Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L. M., Nishikawa, S., Chiba, T., Era, T., and Nishikawa, S. (2005) Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat. Biotechnol.* **23**, 1542–1550

16. D’Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E., and Baetge, E. E. (2005) Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* **23**, 1534–1541

17. Katsumoto, K., and Kume, S. (2011) Endoderm and mesoderm reciprocal signaling mediated by CXCL12 and CXCR4 regulates the migration of angioblasts and establishes the pancreatic fate. *Development* **138**, 1947–1955

18. Haab, B. B. (2005) Antibody arrays in cancer research. *Mol. Cell Proteomics* **4**, 377–383

19. Yalcin, A., Clem, B. F., Simmons, A., Lane, A., Nelson, K., Clem, A. L., Brock, E., Slow, D., Wattenberg, B., Telang, S., and Chesney, J. (2009) Nuclear targeting of 6-phosphofructo-2-kinase (PFKFB3) increases proliferation via cyclin-dependent kinases. *J. Biol. Chem.* **284**, 24223–24232

20. Zhang, X., Neganova, I., Przyborski, S., Yang, C., Cooke, M., Atkinson, 1947–1955
S. P., Anyfantis, G., Fenyk, S., Keith, W. N., Hoare, S. F., Hughes, O., Strachan, T., Stojkovic, M., Hinds, P. W., Armstrong, L., and Lako, M. (2009) A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. J. Cell Biol. 184, 67–82

21. Tomizaki, K. Y., Usui, K., and Mihara, H. Protein-protein interactions and selection: array-based techniques for screening disease-associated biomarkers in predictive/early diagnosis. FEBS J. 277, 1996–2005

22. Kusnezow, W., Banzon, V., Schroder, C., Schaal, R., Hoheisel, J. D., Ruffer, S., Luft, P., Duschi, A., and Syagailo, Y. V. (2007) Antibody microarray-based profiling of complex specimens: systematic evaluation of labeling strategies. Proteomics 7, 1786–1799

23. Wingren, C., Ingvarsason, J., Dexlin, L., Szul, D., and Borrebaeck, C. A. (2007) Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support. Proteomics 7, 3055–3065

24. Schroder, C., Jacob, A., Tonack, S., Radon, T. P., Sill, M., Zucknick, M., Ruffer, S., Costello, E., Neoptolemos, J. P., Crnogorac-Jurcevic, T., Bauer, A., Fellenberg, K., and Hoheisel, J. D. Dual-color proteomic profiling of complex samples with a microarray of 810 cancer-related antibodies. Mol. Cell Proteomics 9, 1271–1280

25. Ko, I. K., Kato, K., and Iwata, H. (2003) Parallel analysis of multiple surface markers expressed on rat neural stem cells using antibody microarrays. Biomaterials 26, 4882–4891

26. Wu, J. Q., Wang, B., Belov, L., Dyer, W. B., Zaunders, J., Cunningham, A. L., Dwyer, D. E., and Saksena, N. K. (2007) Antibody microarray analysis of cell surface antigens on CD4+ and CD8+ T cells from HIV+ individuals correlates with disease stages. Retrovirology 4, 83

27. Belov, L., de la Vega, O., dos Remedios, C. G., Mulligan, S. P., and Christopherson, R. I. (2001) Immunophenotyping of leukemias using a cluster of differentiation antibody microarray. Cancer Res. 61, 4483–4489

28. Kaufman, K. L., Belov, L., Huang, P., Mactier, S., Scolyer, R. A., Mann, G. J., and Christopherson, R. I. (2010) An extended antibody microarray for surface profiling metastatic melanoma. J. Immunol. Methods 358, 23–34

29. Cowan, C. A., Klimanskaya, I., McMahon, J., Alienza, J., Witmyer, J., Zucker, J. P., Wang, S., Morton, C. C., McMahon, A. P., Powers, D., and Melton, D. A. (2004) Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. 350, 1353–1356

30. Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J., Agulnick, A. D., D’Amour, K. A., Carpenter, M. K., and Baetge, E. E. (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. Nat. Biotechnol. 26, 443–452

31. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Bioinformatics 4, 249–264

32. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57

33. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13

34. Chen, D. S., Soen, Y., Stuge, T. B., Lee, P. P., Weber, J. S., Brown, P. O., and Davis, M. M. (2005) Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. PLoS Med. 2, e265

35. Soen, Y., Chen, D. S., Kraft, D. L., Davis, M. M., and Brown, P. O. (2003) Detection and characterization of cellular immune responses using peptide-MHC microarrays. PLoS Biol. 1, E65