Nonlinear electrophoretic response yields a unique parameter for separation of biomolecules

Joel Pel[a], David Broemeling[a], Laura Mai[a], Hau-Ling Poon[b], Giorgia Tropini[a], René L. Warren[c], Robert A. Holt[c], and Andre Marziali[a,1]

[a]Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada; [b]Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada; and [c]Michael Smith Genome Sciences Centre, Vancouver, BC V5Z 1L3, Canada

Communicated by Ronald W. Davis, Stanford University School of Medicine, Stanford, CA, July 7, 2009 (received for review February 23, 2009)

We demonstrate a unique parameter for biomolecule separation that results from the nonlinear response of long, charged polymers to electrophoretic fields and apply it to extraction and concentration of nucleic acids from samples that perform poorly under conventional methods. Our method is based on superposition of synchronous, time-varying electrophoretic fields, which can generate net drift of charged molecules even when the time-averaged molecule displacement generated by each field individually is zero. Such drift can only occur for molecules, such as DNA, whose motive response to electrophoretic fields is nonlinear. Consequently, we are able to concentrate DNA while rejecting high concentrations of contaminants. We demonstrate one application of this method by extracting DNA from challenging samples originating in the Athabasca oil sands.

concentration | DNA | electrophoresis | purification | SCODA

Methods for separating different molecular species are the cornerstone of analytic techniques in molecular biology. Of these, nucleic acid extraction from complex sources is a molecular separation problem of great importance to current challenges in genomics, metagenomics, forensics, biodefense, food and water safety, and clinical molecular diagnostics.

The ubiquitous column- and bead-based nucleic acid extraction methods that dominate the field of nucleic acid extraction use selective chemical affinity between nucleic acids and ion exchange or similar resins and beads to capture target molecules. Although these methods often involve mechanical steps including filtration and centrifugation, they are relatively inexpensive and work well in a variety of samples. Their inadequacy lies in the fact that the separations are based on chemical affinity and therefore perform poorly in the presence of contaminant molecules that either have similar chemical properties to nucleic acids or foul the capture matrix (1). Precipitation methods are often used after column or bead extractions to remove contaminants that carry through; however, this further reduces yield of the methods, particularly in cases with low target concentrations.

This weakness of existing methods is a critical problem for DNA extraction from environmental samples. For example, humic acids, a family of contaminants abundant in soil, coextract with DNA in phenol-based separations owing to their solubility in the aqueous phase (1) and partly carry through column- and bead-based methods. The situation worsens with a low starting concentration of nucleic acids because the dual challenge must then be faced of concentrating few nucleic acids while rejecting large amounts of contaminants. A universal, simple, and highly selective method to concentrate DNA from contaminated and low-abundance sources would be very desirable.

We have found a unique solution to this problem in the physics of electrophoresis. It has long been known that nucleic acid molecules, because of their exceptionally long contour lengths and high linear charge density, exhibit complex electrophoretic behavior when reptating through a separation medium such as agarose gel (2). In some electric field regimens, this can include a highly nonlinear response of the drift velocity to changes in field magnitude (3). We previously showed that, by exploiting this nonlinear response, we could induce net drift of DNA molecules under the influence of 2 synchronously rotating electric fields, each of which has a zero time-averaged magnitude and would individually impart no net drift to the molecules (4). Failure of the superposition principle, and the resulting net drift, occurs only for molecules whose electrophoretic response is nonlinear. Consequently, molecules with highly nonlinear response can be selected for drift over other ions and biomolecules.

Two significant realizations arose from this work. First, nonlinear electrophoretic response could be used as a physical parameter for separating biomolecules. Second, the velocity field pattern generated with this method could be made to diverge in regions of the separation medium that did not contain current sources or sinks. Maxwell's equations governing electric current in conducting media specify that current fields must be free of divergence except at electrodes, making DNA concentration based on static electric fields impractical because of electrochemical damage of the DNA. The velocity field generated by our method concentrates biomolecules in regions free of electrodes.

We use a combination of rotating dipole and quadrupole electric fields in an aqueous gel (Fig. 1), in a method termed *synchronous coefficient of drag alteration* (SCODA) (4, 5), for generating a divergent velocity field that is capable of selectively concentrating nucleic acids in a gel. A brief explanation of the method follows.

Electrophoretic mobility μ is defined as $\vec{v} = \mu\vec{E}$, where v is the velocity and μ is typically considered constant with respect to electric field E. Although mobility is approximately constant for most molecules, it is not so for reptating nucleic acids. The contour length of a nucleic acid molecule is typically many persistence lengths, leading to a large amount of conformational entropy, which when coupled with the field sensitivity induced by the molecule's large linear charge density leads to a strong dependence of its mobility on field strength (6). A more accurate expression for the mobility of a nucleic acid is therefore $\mu(E) = \mu_0 + kE$, where E is the magnitude of the field and k is the linear field dependence of the mobility, which captures the quadratic dependence of reptating DNA velocity on field. We exploit this quadratic behavior, by applying a driving field rotating at frequency ω_1, which by virtue of the quadratic dependence creates a frequency-doubled component at $2\omega_1$. We then heterodyne this double-frequency component with a mobility-modulating quadrupole field at $2\omega_1$, leading to a constant drift-velocity term that inherits the radial geometry of the quadrupole field. The result is an average drift velocity that is proportional to kE_dE_qr, where E_d and E_q are the dipole and quadrupole

Author contributions: J.P., D.B., G.T., R.A.H., and A.M. designed research; J.P., D.B., L.M., H.-L.P., G.T., and R.L.W. performed research; J.P., D.B., G.T., R.A.H., and A.M. analyzed data; and J.P., R.A.H., and A.M. wrote the paper.

Conflict of interest statement: J.P., D.B., H.-L.P., and A.M. have a financial interest in Boreal Genomics, exclusive licensee of the technology presented in this paper.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: andre@physics.ubc.ca.

This article contains supporting information online at www.pnas.org/cgi/content/full/0907402106/DCSupplemental.

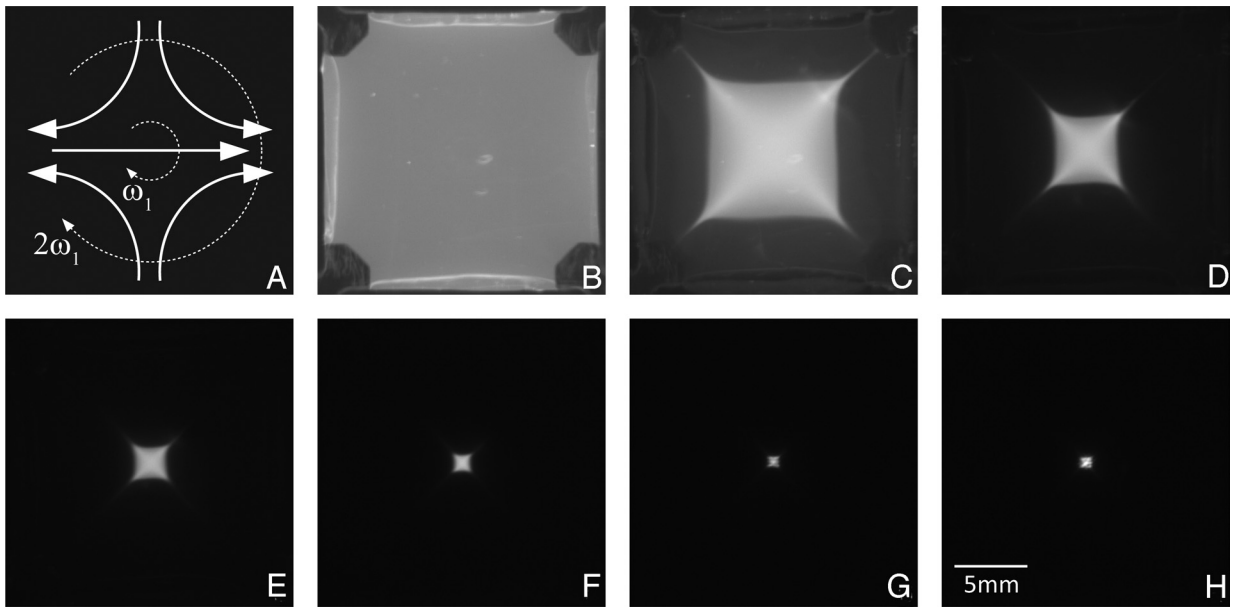

Fig. 1. SCODA concentration sequence. Time-lapse sequence showing concentration of SYBR Green I–stained pUC19 DNA (2.7 kb) from a homogeneous solution of 0.2 ng/μL of DNA in 1% agarose and 0.25× TBE, to a 750-μm-diameter spot. Images are taken at 10-min intervals, for a total run time of 60 min at a SCODA field of 250 V/cm (maximum field in SCODA gel). The concentration of DNA in the focused spot is estimated to be 100–200 ng/μL. (*A*) Diagram of dipole and quadrupole SCODA field lines. (*B–H*) SCODA duration, in minutes: $B = 0$, $C = 10$, $D = 20$, $E = 30$, $F = 40$, $G = 50$, and $H = 60$. Camera exposure is reduced to avoid saturation from increasing fluorescence intensity over the course of concentration. Exposure times, in milliseconds: $B = 1{,}000$, $C = 750$, $D = 250$, $E = 50$, $F = 50$, $G = 10$, and $H = 10$.

field amplitudes, respectively, and r is the distance from the center of the field pattern (4). The drift velocity points toward the center of the field pattern for all locations in the gel, thus leading to a divergent velocity field that can be used for molecule concentration.

During application of this field, increasing molecule concentration at the focus is counteracted by diffusion, and the radius of the steady-state molecule distribution at the focus is proportional to $(D/k)^{1/2}$ (4), where D is the molecule's diffusion constant in the gel. The inverse, $(k/D)^{1/2}$, is therefore a unique molecular parameter that will determine whether molecules focus under the effect of these fields. Fortuitously, this parameter is generally much larger for nucleic acids than for other molecules. It can consequently be used to preferentially separate and concentrate nucleic acids from complex solutions.

Results and Discussion

The effect of this velocity field is demonstrated in Fig. 1. Plasmid (pUC19) dsDNA is initially uniformly dispersed in agarose gel stained with SYBR Green I (Applied Biosystems). When the SCODA rotating fields (Fig. 1*A*) are applied, the DNA migrates toward the center of the gel, where it eventually reaches a tight focus (Fig. 1*H*). The radial asymmetry in the pattern formed during the first minutes of focusing results from the fact that the fields applied to the gel are discrete approximations (both in spatial geometry and in time) of the ideal dipole and quadrupole fields (4).

The concentration ratio achievable from the implementation shown in Fig. 1 is limited by the size of the gel relative to the size of the focus. The latter is a function of molecular parameters and is independent of gel dimensions. Consequently, large concentration factors could be achieved with a large gel, although such an implementation has the drawbacks of being unwieldy and requiring the sample to be cast in a gel. To avoid both these deficiencies, we couple electrokinetic injection of the sample with a small (1.5 cm × 1.5 cm) SCODA concentration gel.

In the simplest implementation of electrokinetic injection, an aqueous sample is placed in a reservoir adjoining one side of the SCODA gel (Fig. 2), and a direct current (DC) electric field is applied across the sample and SCODA gel to drive negatively charged molecules into the gel. Once injection is complete, rotating SCODA fields are applied to the gel, focusing molecules with large values of k/D, while moving contaminant molecules with negligible k/D values in circles. Contaminants can be electrophoretically washed out of the gel by superimposing a small DC bias field on the SCODA fields. This causes low-k/D molecules to be pushed off the edge of the gel because they are not contained by the SCODA force. The injection method itself is therefore doubly selective: by applying an injection field of the correct sense, only negatively charged molecules are injected into the gel; of these, molecules with large k/D are trapped and focused, whereas low-k/D molecules drift through the gel into the anode buffer chamber.

Because there is no fluid flow, particulates such as sand and soil particles can remain in the sample chamber, eliminating a fractionation step required in many alternate purification techniques and reducing the potential for loss of valuable DNA. Although in regular electrophoretic separation, such particles could adversely affect dispersion of the sample band during injection, such dispersion is inherently counteracted during SCODA focusing and therefore does not affect the performance of the concentration and separation. Furthermore, contaminants transiting through the gel have not been found to interfere with concentration and are eventually cleared from the gel.

Although injection and concentration can be performed simultaneously, stacking of DNA as it enters the gel during injection allows the option of injecting briefly at high voltage until the sample chamber is depleted of DNA and then concentrating the DNA in the gel. Fig. 3 is a demonstration of injection of pUC19 DNA followed by concentration. It is noticeable from these images that the linear displacement of DNA in the gel during the injection time is less than the distance from the focus location to the edge of the gel. Consequently, if large sample volumes and large concentration factors are required, one can deplete and concentrate DNA from one 5-mL sample, then replace the sample with a second 5-mL sample and repeat the injection and concentration process to accumulate more DNA into the focus region. With this method,

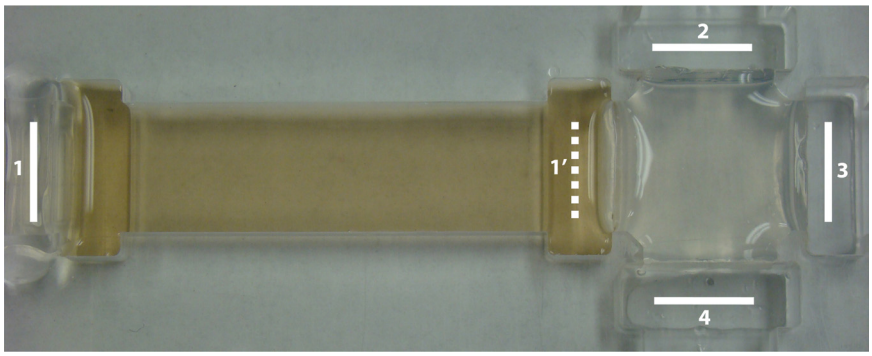

Fig. 2. SCODA gel and injection chamber. SCODA gel boat for electrokinetic injection and concentration shown with a 60-μg/mL humic acid sample in the injection chamber. As indicated in the overlay, electrodes placed at locations 1 and 3 allow for application of a DC electric field to inject negative ions into the gel, where the rotating SCODA fields (applied at electrodes 1′, 2, 3, and 4) concentrate and trap molecules with a high ratio of *k/D,* whereas low-*k/D* molecules do not concentrate. This allows selective trapping and concentration of nucleic acids in the center of the gel. 1′ is a high-impedance electrode only used to monitor and clamp potentials; current is sourced at 1. Opposite electrodes (1′ and 3, 2 and 4) surrounding the SCODA gel are spaced 25 mm apart.

concentration factors are limited only by the process time, not by the gel geometry, and factors >10,000 have been demonstrated.

Using this method, we have concentrated dsDNA ranging in size from 200 bp to 1.6 Mbp; the absence of fluid flow allows concentration of fragments >1 Mbp without shearing. We have also demonstrated direct extraction from samples containing particulates by adding soil directly to our sample chamber, lysing in situ, and concentrating to yield intact nucleic acids of lengths in excess of 600 kb. Finally, we have shown concentration of RNA and denatured proteins under appropriate conditions, although this will not be discussed further here.

In addition, the field rotation frequency can be used to reject long DNA fragments if desired, by rotating the fields faster than the relaxation time of the fragments in the gel. At moderate fields (≈50 V/cm) in 4% agarose gels made with 0.25× Tris/Borate/EDTA (TBE) buffer, slow field rotations (4 s) will concentrate all fragment lengths up to 50 kb, whereas fields rotating at 160 ms will only focus fragments up to 3 kb in length. By adjusting the phase of the fields, defocusing of fragments can also be achieved, allowing a variety of low-pass, high-pass, and band-pass schemes to be applied for DNA length selection.

The physical nature of the concentration process leads to some significant advantages, one of which is that concentration and selectivity do not require reversible binding to a matrix. Such binding in other methods is not always very efficient or fully reversible, leading to low concentration efficiencies, particularly at low molecule concentrations. The efficiency of SCODA concentration, as measured using quantitative real-time (QRT) PCR, is 91% with an SD of 28% for extraction from 56 fM (2×10^{-4} ng/μL) DNA in buffer. The concentration enhancement factor from a single 5-mL injection to a 500-μm-diameter focused spot of DNA is ≈4,000.

SCODA concentration efficiency remains high when working with very few target molecules. We tested SCODA's ability to recover DNA from 1.72×10^8 molecules on input (56 fM when diluted to 5 mL) down to 172 molecules (56 zM when diluted to 5 mL, or ≈35 molecules per milliliter). In these experiments, serial dilutions of ΦX174 dsDNA in buffer were diluted to 5 mL and applied to the SCODA sample chamber, injected and concentrated as described above, then extracted in ≈20-μL gel cores and analyzed using QRT-PCR to measure successful amplification. One hundred percent of the purified samples, from 1.72×10^8 molecules down to 172 molecules input (56 zM), amplified successfully.

Excellent rejection of contaminants such as humic acids can be achieved during the concentration and extraction process. This is done by emptying the sample chamber once injection is completed, replacing the sample with running buffer, and continuing the focusing process with a small DC bias field applied in the opposite direction of the injection field. The DC bias electrophoretically pushes low-*k/D* molecules that coinjected with the DNA back out of the gel to the anode buffer chamber, resulting in clearance of humic acids while the DNA is held in the center of the gel by the concentration fields (Fig. 4).

Using this DC-bias technique we have quantitatively demonstrated SCODA's ability to reject humic acids, a contaminant known to inhibit PCR and other molecular biologic methods (1), by comparing SCODA's performance with that of existing technologies (Fig. 5). The success of humic acid rejection was assayed by the relative success of PCR amplification of humic acid–spiked DNA samples. All samples contained a fixed amount of DNA (2 ng) with increasing mass of humic acid, as indicated in Fig. 5. DNA extraction performance is therefore compared on the basis of mass ratio of DNA to humic acids, to account for the different input volumes of the other methods and the possibility of dilution or concentration before sample extraction. Input volumes for the silica column, magnetic bead, and SCODA methods are 100 μL, 25 μL, and 5 mL, respectively.

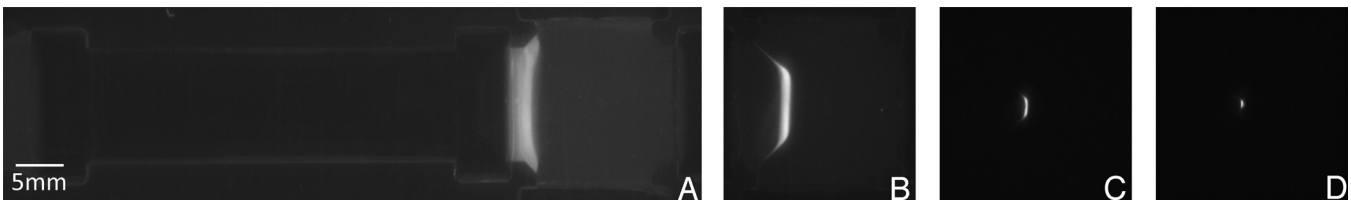

Fig. 3. Injection and concentration sequence. Time-lapse sequence demonstrating injection and concentration of 200 ng of SYBR Green 1–stained pUC19 DNA. DNA is injected from 5 mL of 0.05× TBE buffer into a 1% agarose gel made with 0.25× TBE buffer. (*A*) Image taken after 10 min of injection at 20 V/cm. (*B*) Image taken after 10 min of subsequent SCODA with a maximum field of 250 V/cm. (*C* and *D*) Images taken at incremental 20-min SCODA concentration intervals for a total run time of 60 min. Camera exposure, in milliseconds: $A = 1{,}000$, $B = 500$, $C = 100$, and $D = 20$.

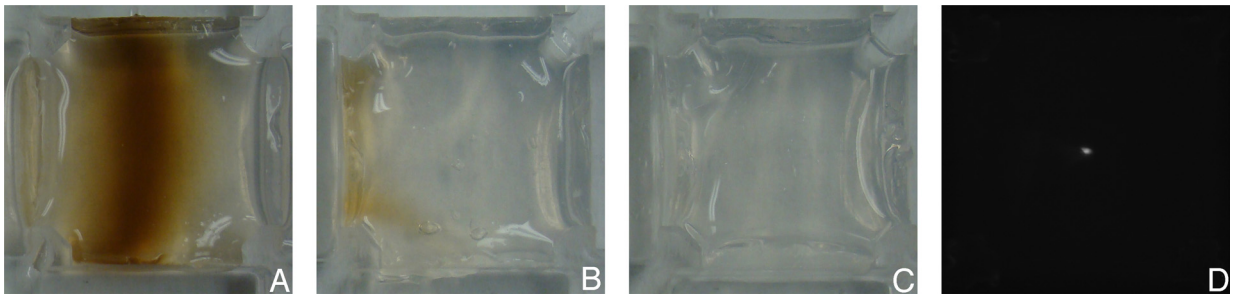

Fig. 4. Electrophoretic washing. Time-lapse sequence showing dispersion and removal of contaminants concurrently with concentration of the desired nucleic acids. Images are a continuation of the experiment shown in Fig. 2, where 200 ng of pUC19 DNA was spiked into a 60-μg/mL humic acid solution, injected, and concentrated with superimposed electrophoretic washing. The increasing clarity of the gel indicates that the negatively charged humic acid contaminants are migrating out of the gel under the applied fields. (*A*–*C*) Visible light images showing decreasing humic acid (brown stain) contamination of the gel: in *A* after 20 min of DC injection, in *B* after an additional 10 min of SCODA concentration with DC field applied to wash contaminants from the gel, and in *C* after an additional 50 min of concentration with DC field. (*D*) UV-transilluminated image (100-ms exposure) taken at the same time point as *C* (80 min total elapsed time), in which stained DNA is clearly visible in the center of the gel.

SCODA rejection of humic acids is excellent, and although other preparation methods exist that are specifically designed to remove humic acids, the SCODA method is not specific to this contaminant and has been shown to behave similarly with other contaminants.

Similar contaminant rejection tests were carried out with a combination of proteins from lysed *Escherichia coli* cell culture and BSA, indicating again that affinity-based methods are more easily overwhelmed by contaminants. By comparison with a spin column (Qiagen QIAprep Spin Miniprep Kit), SCODA was more than 60-fold more effective at rejecting a mixture of *E. coli* protein contaminants and BSA and did not suffer substantial reduction in DNA yield as the amount of input contaminant was increased (Table S1).

The apparatus presented is limited to the application of 90 V at each of the 4 electrodes, and consequently concentration requires at least 1 h. A prototype instrument in development is capable of 1,000-V potentials at the source electrodes and has initially been used to demonstrate concentration of DNA fragments 2.7 kb in length from 1 mL of sample in as little as 6 min.

DNA Extraction from Athabasca Oil Sands. SCODA's ability to reject contaminants including humic acids, its high efficiency in samples with low DNA abundances, and its ability to work directly with samples containing particulates make it an ideal method for environmental DNA extraction. Although the majority of biological diversity in our environment remains unexplored, a number of metagenomics projects, or surveys of such diversity through DNA sequencing, are under way or have been carried out (7–9).

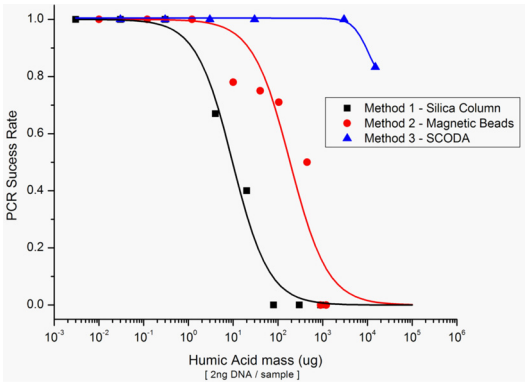

Fig. 5. Humic acid contaminant rejection. Success of PCR amplification of 2-ng samples of ABI Quantifiler Human DNA Standard containing increasing humic acid mass (horizontal axis, logarithmic scale), after extraction with silica column, magnetic bead, and SCODA purification methods. The 50% cutoff for SCODA is beyond the range of humic acid concentrations that could be mixed from commercially available humic acid stocks. Projected performance of SCODA in humic acid rejection is at least 100-fold better than the next commercially available method that is not specifically designed to reject humic acids.

We applied SCODA to a particularly difficult metagenomics project that had hitherto been impeded by failure of DNA extraction: sampling the biological diversity of subsurface material from the Athabasca oil sands. Defining the microbial flora associated with petroleum reservoirs will facilitate studies of petroleum microbiology, and resident microbes may also provide a source of novel enzymes for industrial applications. The Athabasca oil sands comprise a mixture of bitumen (heavy oil), water, sand, and clay. Quartz sand forms the bulk of the material, and it is thought that individual grains of sand are covered in a film of water that is in turn surrounded by bitumen, such that the oil fraction is not in direct contact with the mineral grains (10).

For the present study, subsurface material taken from a drilling core (Fig. 6) was prepared by resuspension in buffer and removal of the oil layer. The remaining unfractionated sand/silt and buffer sample was subjected to a cell lysis protocol, and the entire solution was transferred to the SCODA injection chamber in multiple 5-mL volumes for repeated injection and concentration of DNA. In subsequent experiments with 10 g and 50 g of raw oil sand drilling core material, agarose plugs containing the concentrated DNA were successfully extracted and estimated, to an order of magnitude based on DC gel fluorescence intensity of controls, to yield $\approx$100 ng and $\approx$1 μg, respectively.

Fosmid sequencing libraries were constructed using the DNA supplied, with no additional clean-up or purification other than agarase extraction from the agarose plugs. A total of 1,124 sequence reads were generated. Characterization of oil sands–associated environmental DNA was performed by searching the largest public repositories of characterized DNA and peptide sequences, GenBank-nt and GenBank-nr, respectively. Best-scoring nucleotide Blastn (11) hits to GenBank-nt were used to determine similarity of sampled genetic material to that of known organisms. The sample seems diverse, with matches to >200 distinct bacterial genomes (Fig. 7). With caution, given that these data are from a single sample, it seems that this observation is consistent with the presence of microbial flora in the subsurface oil sand environment. It is also important to note that because drilling conditions are nonsterile, we cannot exclude the possibility that there is representation of surface organisms in the sequence data. By Blastn, the largest number of best-scoring matches were obtained against the genomes of aromatic hydrocarbons-degraders *Novosphingobium aromaticivorans* and *Rhodopseudomonas palustris* (12), metal-reducer *Anaeromyxobacter dehalogenans* (13), secondary metabolite producer *Streptomyces avermitillis* (14), and to well-characterized soil-dwelling acti-

Fig. 6. Oil sands. *Left:* raw oil sand sample, showing bitumen, water, sand, and clay mixture. *Right:* Resuspension of oil sands mixture in buffer for DNA extraction from SCODA.

nomycetes such as *Nocardia farcinica* (15) and *Rhodococcus sp.* RHA1 (16, 17). The G + C content of the fosmid clone sequences ranged from 24% to 67%, and Blastn matches generally show sequence conservation between 30% and 70% with a Gaussian-like distribution (Fig. 7) and an average match length of 705 ± 33 nt. Together, these results show that the SCODA method yields DNA suitable for library construction and metagenomic analysis and suggest that genetic material associated with the oil sand environment is from divergent organisms that have not been identified or characterized previously.

Materials and Methods

As illustrated in Fig. 2, the sample is placed in a 5-mL chamber between the focusing gel and a rear gel dam. Electrodes 1 and 3 are used to establish a DC electrokinetic injection field, with electrode 1 as the cathode. Superimposed or alternated with the DC field are the SCODA fields, as described in the main text. To arrange the SCODA fields, it is normally necessary to arrange electrodes at locations 1′, 2, 3, and 4 symmetrically around the gel (4). During injection though, electrode 1′ must not source or sink current to avoid damaging the DNA in its proximity. To avoid this, 1′ is used as a virtual electrode to clamp the voltage at that location, whereas electrode 1 is used to source or sink the required current and achieve the desired potential at location 1′, generating the desired electric fields.

Typically, DC injection fields of 20 V/cm and SCODA fields of ≈250 V/cm (maximum value in the gel) are applied, leading to combined injection and focusing times of 1 to 2 h (for molecules between 200 bp and 50 kb).

To extract DNA directly from soil and avoid shearing of long molecules, 0.1 g of soil was cast mixed with 0.4% agarose gel and 0.25× TBE in the 5-mL SCODA sample chamber and incubated for 1 h at 37 °C in the presence of 5 mg/mL of lysozyme and 0.2 mg/mL proteinase K. SDS (1%) was then added and incubated at 65 °C for 30 min. After lysis, DNA was injected into a 1% agarose gel by 10-V/cm fields for 40 min and subsequently concentrated with a maximum field amplitude of 24 V/cm and a field rotation period of 1,440 s to recover high-molecular-weight DNA. After 17 h of concentration (the long run length is due to the low mobility of large DNA fragments in an agarose gel), the agarose plug extracted from the SYBR Green I–stained focus was analyzed by pulsed-field gel electrophoresis and was found to contain DNA fragments up to 600 kb in length. This is an order of magnitude greater in length than with conventional direct extraction techniques from soil, which are typically limited to <50-kb fragments because of shearing (18).

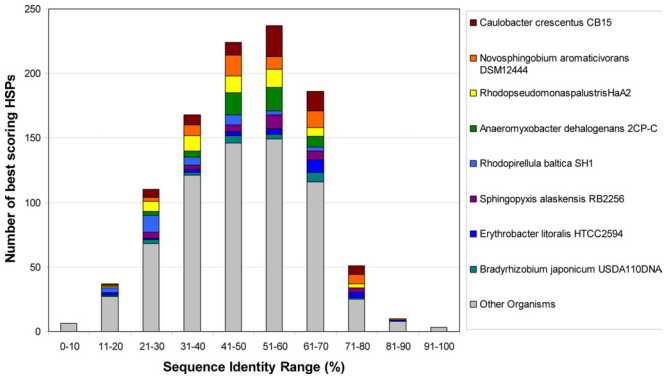

Fig. 7. Conservation of oil sand–associated environmental DNA. Vector and quality trimmed fosmid sequences prepared from SCODA-concentrated DNA were aligned to GENBANK-nt using wuBLAST. The percentage sequence identity distribution for high-scoring pairs (HSPs) follows a Gaussian distribution. The bulk of the fosmid sequences have restricted sequence identity (≈30–70%) compared with known genomes. A partial legend for best scoring hits is shown; for a complete list, see Table S2.

The efficiency of SCODA was measured using QRT-PCR on purified samples. One nanogram of ΦX174 dsDNA was diluted in 5 mL of 0.05× TBE (56 fM) and applied to the SCODA injection chamber. The DNA was injected at 20 V/cm for 45 min into a 1% low-melting-point agarose gel made with 0.25× TBE and concentrated with a 4-s rotational period at 250-V/cm maximum field for 2.5 h without stain to a predetermined focus location, where it was removed in a ≈20-μL agarose plug. The agarose plug was melted and diluted to 3 times its volume with dH_20 before being divided into 5-μL aliquots and inserted (in duplicate) into a QRT-PCR using TaqMan (Applied Biosystems, part no. 4304437) chemistry. Tenfold and 100-fold dilutions of the SCODA sample were also quantified in duplicate. SCODA-purified samples were compared with a standard curve and aliquots of the SCODA input sample, from which efficiency was calculated in 3 separate runs. Variation in SCODA efficiencies comes in part from variation in quantification from the QRT-PCR assay.

Recovery of DNA from low-concentration samples using SCODA was measured by assaying for relative PCR success. Decreasing numbers of molecules of ΦX174 dsDNA in 100 μL of Tris–EDTA (TE) buffer were applied to the SCODA input chamber and diluted to 5 mL with dH_20. SCODA samples were processed as above in the efficiency experiments, after which they were removed in a ≈20-μL agarose plug. The agarose plug was melted and diluted to 3 times its volume with dH_20 before being divided into 5-μL aliquots and inserted (in duplicate) into a QRT-PCR using TaqMan chemistry. Tenfold and 100-fold dilutions of the SCODA sample were also analyzed in duplicate. SCODA-purified samples were compared with a standard curve and aliquots of the SCODA input sample, from which successful amplification was determined.

To test for humic acid rejection, we spiked dH_20 samples containing 2 ng of DNA (Applied Biosystems, Quantifiler Human DNA Standard, part no. 4343895) with increasing masses (see Fig. 5) of humic acids (Sigma-Aldrich, part no. H16752) and purified these samples with Qiagen (QIAquick PCR Purification Kit, part no. 28104), Promega (DNAiq DNA Isolation System, part no. TB297, using the manufacturer's DNA Isolation from Liquid Blood protocol), and SCODA technologies. Qiagen and Promega samples were processed as per the manufacturers' instructions. One hundred microliters of sample was applied to the Qiagen columns, and extracted DNA was eluted in 30 μL of elution buffer. Twenty-five microliters of sample was input to the Promega system, with the final elution volume being 100 μL. The SCODA samples were injected from 5 mL at 10 V/cm for 20 min into a 1% low-melting-point agarose gel made with 0.25× TBE. A static 0.5-V/cm washing field was applied toward the injection chamber to wash contaminants out of the gel as concentration proceeded at 125-V/cm maximum fields and 4-s rotational period for 3 h. A final 2 h of concentration at the same conditions was done without the washing field, so that the DNA was concentrated to a predetermined focus location, where it was removed in a ≈20-μL agarose plug. The agarose plug was melted and diluted to 3 times its volume with dH_20 before being aliquoted

and 5 μL inserted (in duplicate) into a QRT-PCR using SYBR Green (Applied Biosystems, part no. 4309155) chemistry. The success of the humic acid rejection was assayed by the relative success of PCR amplification of the spiked DNA over a number of samples. Two experiments were done for the lowest humic acid input to Qiagen, and 3 experiments for the next 2 humic acid inputs. For the 4-μg and 20-μg inputs, 6 and 15 experiments were done, respectively. The highest 3 inputs to Qiagen were done in triplicate. The lowest 3 humic inputs to Promega were done in triplicate. The 10-, 40-, 105-, and 450-μg inputs were repeated 9, 8, 7, and 4 times, respectively. The final 2 samples were done in duplicate. For the SCODA samples, 1 experiment was done at each humic acid data point, except for the highest humic input, for which 12 experiments were conducted. These experiments were done before there was sufficient cooling on our apparatus to tolerate higher fields, leading to longer run times.

As a general test of contaminant rejection, we concentrated plasmid DNA from a lysed *E. coli* cell culture, then assayed for the presence of proteins in the extracted sample using the bicinchoninic acid (BCA) total protein assay (19) as per the manufacturer's instructions. As a benchmark, the SCODA results were compared with a Qiagen QIAprep Spin Miniprep Kit designed for *E. coli* lysis and plasmid purification. A colony of DH10BT1 *E. coli* cells transformed with pUC19 plasmid containing the ampicillin resistance gene was grown until the OD of the cells at 600 nm was 0.6, and then used for these experiments. Initial results from the Qiagen columns show near-zero protein carry-through using the Qiagen lysis and purification protocol; protein concentration was measured with the BCA assay after lysis (Qiagen protocol step 3) and after purification. To test the protein rejection limits of the methods, 43.75 mg of BSA protein was added to the *E. coli* lysate, which, after purification, showed significant protein carry-through (Table S1). Plasmid DNA recovered in control and spiked samples was verified for length and sequence in a separate purification run by DC gel and QRT-PCR.

To prepare the *E. coli* for SCODA purification, a different lysis method than that of Qiagen was used to avoid the high conductivity associated with the Qiagen lysis buffers, which can be undesirable for high-performance SCODA runs. For lysis in SCODA, 1.5 mL of *E. coli* sample was added to 350 μL of STET buffer containing 250 μg of lysozyme and then boiled for 40 s. The sample was then diluted to 5 mL with dH_20 in the SCODA injection chamber. Injection was performed at 18 V/cm for 40 min into a 1% low-melting-point agarose gel made with sodium borate buffer (4 mM NaOH and 20 mM boric acid). Concentration and contaminant rejection was accomplished by applying a static 2-V/cm washing field toward the injection chamber and concentrating with 125-V/cm maximum SCODA fields for 2 h with a 4-s rotational period. A final 2 h of concentration was done with the same conditions but no DC field applied, so that the DNA was concentrated to a predetermined focus location, where it was removed in a $\approx$20-μL agarose plug, diluted to 100 μL with sodium borate buffer, and analyzed with the BCA assay. Because of the addition of lysozyme in the SCODA lysis, the initial protein concentration of the SCODA samples is higher than that of Qiagen, but comparable final protein concentration is observed after SCODA and Qiagen purification methods. When the SCODA lysate is spiked with BSA as in the Qiagen samples and run with the same rejection conditions as the control sample, minor protein carry-through in SCODA is observed. Again, plasmid DNA recovered in each purification method was verified in a separate run by DC gel and RT-PCR.

Oil sands samples were resuspended in 0.5$\times$ TE buffer (2 mL buffer for every 1 g of oil sands) before manual stirring and pulverization. The separated oil layer was removed and lysozyme added to the remaining sample at a concentration of 150 ng/mL, then shaken at 37 °C for 2 h. Triton X-100 was added at 5% vol/vol, and the sample was vortexed and subjected to 3 freeze/thaw cycles. Finally, proteinase K was added to a concentration of 100 μg/mL, and the solution was placed in a 50 °C water bath for 1 h. Five-milliliter aliquots were loaded into the SCODA injection chamber for subsequent injection and concentration. Fluorescent controls used for DNA quantification indicate DNA concentration factors >10,000 from these experiments, arising from multiple injections.

Samples were injected at 10 V/cm for 15 min into a 1% low-melting-point agarose gel made with 0.25$\times$ TBE. After sample injections, the sample was removed from the sample chamber and replaced with fresh buffer during concentration. Concentration proceeded with 125-V/cm fields and 4-s rotational period for 4 h to a precalibrated focus location, where the DNA was extracted in a $\approx$50-μL agarose plug before agarase digestion.

SCODA-concentrated DNA fragments from the oil sands were ligated directly to pEpiFos5 vector without size selection, then packaged and plated using the EpiFOS fosmid library production kit (Epicentre, catalog no. FOS0901) according to the manufacturer's instructions. Fosmid end sequences were obtained using pEpiFos5-Forward and pEpiFos5-Reverse primers in 4-μL reactions containing 0.33 μL BigDye terminators v.3.1 (Applied Biosystems) and 2 μL of $\approx$0.5 μg/μL of alkaline lysis–purified fosmid DNA. The sequences were read using a 96-capillary 3730*xl* DNA Analyzer (Applied Biosystems). A total of 1,152 total fosmid sequences were obtained, and these were vector trimmed using cross-match (20) and quality trimmed using trim2 (-M 10) (21). The resulting 1,124 sequences were aligned to GenBank-nt and GenBank-nr using wuBLASTn (BLAST version 2.0) and wuBLASTx (BLAST version 2.0), respectively. The default parameters were used for both programs, and only the best-scoring match from each fosmid read was subsequently evaluated.

We also used Blastx to search the GenBank-nr database with predicted amino acid sequences from the fosmid end read.

Summary. The successful sequencing of DNA extracted from the Athabasca samples would not have been easily accomplished with conventional methods, owing to the high level of humic substance contamination from the oil. In addition, the DNA yield achieved through successive loads of the SCODA sample chamber was sufficient to allow library construction without prior whole-genome amplification, which is known to generate sequence representation bias.

This work demonstrates the utility and selectivity of nonlinear electrophoretic response for biomolecule separations. Although the operating parameters demonstrated in this article are optimized for nucleic acids, many applications remain to be explored, and it is expected that the proof-of-concept demonstrated in this article will lead to other applications for nucleic acids and possibly other biomolecules. The purely electrophoretic nature of the SCODA separation method yields additional benefits, including concurrent size selection, insensitivity to sample contaminants and debris, and the ability to maintain high molecular weight of the product. We expect this new tool to find broad application in genomics, metagenomics, food and water safety, environmental DNA detection, and similar applications in which starting sample materials are complex, dilute, and presently difficult to analyze.

ACKNOWLEDGMENTS. We thank Dr. Bruce C. W. McGee (E-T Energy, Calgary, Alberta, Canada) for providing the oil sands samples used in this work; and Jason R. Dwyer for his insightful comments on and editing of the manuscript. This project was supported by Grant R01HG003640 from the National Human Genome Research Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Human Genome Research Institute or the National Institutes of Health. Part of this work was also supported by Grant PPP-77272 from the Canadian Institutes of Health Research.

1. Rajendhran J, Gunasekaran P (2008) Strategies for accessing soil metagenome for desired applications. *Biotechnol Adv* 26:576–590.
2. Viovy JL (2000) Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms. *Rev Mod Phys* 72:813–872.
3. Frumin LL, Peltek SE, Ziberstein GV (2001) Nonlinear electrophoresis and focusing of macromolecules. *J Biochem Biophy Methods* 48:269–282.
4. Marziali A, Pel J, Bizzotto D, Whitehead LA (2005) Novel electrophoresis mechanism based on synchronous alternating drag perturbation. *Electrophoresis* 26:82–90.
5. Broemeling D, et al. (2008) An instrument for automated purification of nucleic acids from contaminated forensic samples. *J Assoc Lab Automation* 13:40–48.
6. Rousseau J, Drouin G, Slater GW (2000) Gel electrophoretic mobility of single-stranded DNA: The two reptation field-dependent factors. *Electrophoresis* 21:1464–1470.
7. Daniel R (2005) The metagenomics of soil. *Nat Rev Microbiol* 3:470–478.
8. Lorenz P, Eck J (2005) Metagenomics and industrial applications. *Nat Rev Microbiol* 3:510–516.
9. Langer M, et al. (2006) Metagenomics: An inexhaustible access to nature's diversity. *Biotechnol J* 1:815–821.
10. Czarnecki J, Radoev B, Schramm LL, Slavchev R (2005) On the nature of Athabasca oil sands. *Adv Colloid Interface Sci* 114:53–60.
11. Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
12. Larimer FW, et al. (2003) Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas. Nat Biotechnol* 22:55–61.
13. Sanford RA, Cole JR, Tiedje JM (2002) Characterization and description of. *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic myxobacterium *Appl Environ Microbiol* 68:893–900.
14. Omura S, et al. (2001) Genome sequence of an industrial microorganism. *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *Proc Natl Acad Sci USA* 98:12215–12220.
15. Ishikawa J, et al. (2004) The complete genomic sequence of *Nocardia farcinica* IFM 10152. *Proc Natl Acad Sci USA* 101:14925–14930.
16. Warren R, et al. (2004) Functional characterization of a catabolic plasmid from polychlorinated- biphenyl-degrading *Rhodococcus* sp. Strain RHA1. *J Bacteriol* 186:7783–7795.
17. McLeod MP, et al. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci USA* 103:15582–15587.
18. Robe P, Nalin R, Capellano C, Vogel TM, Simonet P (2003) Extraction of DNA from Soil. *Eur J Soil Biol* 39:183–190.
19. Smith PK, et al. (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* 150:76–85.
20. Ewing B, Hillier L, Wendl MC, Green P (1998) Base calling of automated sequencer traces using *Phred*. I. Accuracy and assessment. *Genome Res* 8:175–185.
21. Huang X, Wang J, Aluru S, Yang SP, Hillier L (2003) PCAP: A whole-genome assembly program. *Genome Res* 13:2164–2170.