constant and highly vigorous mixing after 40 hours at high temperature. Wang et al., *RNA* 13(1): 151-159 (2007). Hybridization of longer and more complex probe-target mixes likely take longer to approach equilibrium, even if similar temperature stringency and rigorous mixing can be achieved on non-closed platform. As a result, hybridization-capture requires long incubation to approach uniformity and yield, even with ideal probe design and hybridization conditions.

While hybridization-capture can achieve very high complexity (multiple hundreds of thousands of probe sequences simultaneously in a single capture reaction) and can provide continuous and quantitative coverage of longer stretches of the genome, the workflow is long and arduous. A typical workflow involves many time-sensitive, hands-on steps performed by an experienced operator in order to achieve reproducible and consistent results. Further, coverage uniformity and specificity are challenges, which are usually compensated by probe design, hybridization-solution formulation optimization, hybridization duration, and stringency of post hybridization-capture wash steps. Assays still generally suffer from under or over sampling of GC or AT-rich regions.

What are needed are methods and compositions to overcome the existing challenges of current technologies and reduce the complexity of workflows, decrease the hybridization-capture time, increase the capture complex target regions, and improve the specific and uniform capture of target sequences or regions.

SUMMARY

Described herein are methods and compositions for improved hybridization capture for enriching for target nucleic acid sequences from a population of nucleic acids. The method includes: combining in solution a population of nucleic acids and a complementary probe set; permitting the complementary probe set to hybridize or bind to the target nucleic acid; selectively immobilizing the probe/target nucleic acid complex; exposing the selectively immobilized probe/target nucleic acid complex to a temperature below the aggregate T_m to melt unwanted non-target material; and washing the complex to remove the unbound non-target material, thereby enriching for target nucleic acid.

One ~~embodiment~~example described herein pertains to the ability to reduce the hybridization time of the hybridization capture. One ~~aspect~~example described herein reduces the complexity of hybridization capture workflows. Another ~~aspect~~example described herein increases the capture specificity of complex target regions. Another ~~aspect~~example described herein improves the specific capture of target nucleic acid sequences or regions while minimizing sequence bias.

Another ~~embodiment~~example described herein provides for improved hybridization capture compositions. Another ~~aspect~~example described herein provides for improved hybridization capture buffers.

Another ~~embodiment~~example described herein is a method for hybridization capture. The hybridization capture method is a Hyb-Capture-Melt workflow. In a first step a probe panel is contacted with a nucleic acid sample under conditions to promote hybridization. Optionally, the probe panel may contain a capture moiety, for example, but not limited to biotinylation. The biotinylated probes are complementary to target nucleic acids or nucleic acids to be enriched for. In the first step the goal is to increase the hybridization capture yield. Specificity of the nucleic acid capture is secondary to yield. In a second step the biotinylated probes are immobilized to streptavidin beads. Probes with and without hybridized targets are immobilized. In a third step a melt is performed to dissociate all non-specific hybridization interactions. Without being bound to a single theory it is also expected that after the melt is performed re-hybridization of captured probes with target sequence occurs, further improving both yield and specificity. In a fourth step a wash step is performed to further remove un-hybridized and weakly hybridized non-targeted nucleic acids.

Another ~~embodiment~~example described herein is improved compositions for hybridization capture. In one ~~aspect~~example, hybridization capture buffers are provided that improve the hybridization capture efficiency. In another ~~aspect~~example, T_m modifying reagents may be added. In another ~~aspect~~example, additives that decrease T_m dependence on double stranded target GC base content may be added to improve capture.

Another ~~embodiment~~example described herein is a method for hybridization capture. In one step, a hybridization at low stringency is provided. In another step, capture (immobilization) is performed at a temperature sufficient to promote immobilization. In another step, a melt is performed at an elevated temperature to reduce off-target hybridization. In another step, a wash is performed to reduce off-target hybridization and optionally this wash step may be in the absence of a hot-wash.

One ~~embodiment~~ aspect described herein is a method for enrichment of a population of nucleic acid target sequences in a sample comprising: (a) providing a sample of nucleic acid molecules comprising a plurality of target nucleic acid sequences and a plurality of off-target nucleic acid sequences; (b) hybridizing the sample to a panel of nucleic acid probes that are complementary to the plurality of target nucleic acid sequences under hybridization conditions to generate a probe/target complex; (c) selectively immobilizing the probe/target complex to form an immobilized probe/target complex; (d) heating the immobilized probe/target complex to a

temperature at or below the aggregate T_m of the probe/target complex for a period of time sufficient to disassociate the off-target nucleic acid sequences; and (e) washing the immobilized probe/target complex to remove non-hybridized nucleic acid sequences and off-target nucleic acid sequences from the hybridized plurality of target nucleic acid sequences, thereby enriching for the plurality of target nucleic acid sequences in the sample. In one ~~aspect~~example, the panel of nucleic acid probes that are complementary to the plurality of target nucleic acid sequences further comprises a capture moiety. In another ~~aspect~~example, the capture moiety is biotin. In another ~~aspect~~example, the hybridization conditions comprise a hybridization buffer comprising one or more of: (a) salts selected from one or more of monovalent, divalent, sodium, ammonium, cesium, or manganese, chlorides, citrates, sulfates, perchlorates, isothiocyanates, or guanidinium; (b) chelating agents selected from one or more of ethylenediaminetetraacetic acid (EDTA), ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA), or (1,2-bis(o-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid (BAPTA); (c) buffering agents selected from one or more of tris(hydroxymethyl)aminomethane (Tris); 2-(bis(2-hydroxyethyl)amino)acetic acid) (Bicine); (*N*-[tris(hydroxymethyl)methyl]glycine (Tricine); [tris(hydroxymethyl)methylamino] propanesulfonic acid (TAPS); 3-[*N*-tris(hydroxymethyl)methylamino]-2-hydroxypropanesulfonic acid (TAPSO); 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES); 2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid (TES); 3-(N-morpholino)propanesulfonic acid (MOPS); piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES); 2-(*N*-morpholino)ethanesulfonic acid (MES); or dimethylarsenic acid (Cacodylate); (d) detergents selected from one or more of sodium dodecylsulfate (SDS), polysorbate 20 (Tween® 20), octyl phenol ethoxylate (Triton X-100), octylphenoxypolyethoxyethano (IGEPAL®-CA 630), nonyl phenoxypolyethoxylethanol (NP-40), or cetrimonium bromide (CTAB); or (e) additives selected from one or more of glycine, betaine, glycine-betaine, 7-deaza-2′-deoxyguanosine, dimethyl sulfoxide (DMSO), polyethylene glycols 400–1,000,000, glycerol, magnesium, tetramethyl ammonium chloride (TMAC), tetraethylammonium chloride (TEAC), triethylamine hydrochloride, ethylene carbonate, dextran sulfate, or bovine serum albumin (BSA). In another ~~aspect~~example, the hybridization buffer contains formamide. In another ~~aspect~~example, the hybridization buffer does not contain formamide.

In one ~~embodiment~~example, the hybridization conditions comprise an incubation time ranging from 10 minutes to 48 hours. In one ~~aspect~~example, the hybridization conditions comprise an incubation time ranging from 2 hours to overnight. In another ~~aspect~~example, equivalent specificity and equivalent yield are achieved with a 2-hour incubation time as compared to an overnight incubation time. In another ~~aspect~~example, the hybridization conditions comprise an incubation temperature ranging from

about 55 °C to about 75 °C. In another ~~aspect~~example, the hybridization conditions comprise an incubation temperature ranging from about 60 °C to about 70 °C. In another ~~aspect~~example, the hybridization conditions comprise an incubation time of 2 hours and an incubation temperature of about 65 °C.

In another ~~embodiment~~example, the probe/target complex is selectively immobilized using streptavidin beads. In one ~~aspect~~example, the probe/target complex is selectively immobilized under conditions comprising an incubation time ranging from about 10 minutes to about 48 hours. In another ~~aspect~~example, the probe/target complex is selectively immobilized under conditions comprising an incubation temperature ranging from about 20 °C to about 40 °C. In another ~~aspect~~example, the probe/target complex is selectively immobilized under conditions comprising an incubation time of about 30 minutes and an incubation temperature of room temperature. In another ~~aspect~~example, the heating of the immobilized probe/target complex comprises a melt buffer comprising one or more of: (a) salts selected from one or more of monovalent, divalent, sodium, ammonium, cesium, or manganese, chlorides, citrates, sulfates, perchlorates, isothiocyanates, or guanidinium; (b) chelating agents selected from one or more of ethylenediaminetetraacetic acid (EDTA), ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA), or (1,2-bis(o-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid (BAPTA); (c) buffering agents selected from one or more of tris(hydroxymethyl)aminomethane (Tris); 2-(bis(2-hydroxyethyl)amino)acetic acid) (Bicine); (*N*-[tris(hydroxymethyl)methyl]glycine (Tricine); [tris(hydroxymethyl)methylamino] propanesulfonic acid (TAPS); 3-[*N*-tris(hydroxymethyl)methylamino]-2-hydroxypropanesulfonic acid (TAPSO); 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES); 2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid (TES); 3-(N-morpholino)propanesulfonic acid (MOPS); piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES); 2-(*N*-morpholino)ethanesulfonic acid (MES); or dimethylarsenic acid (Cacodylate); (d) detergents selected from one or more of sodium dodecylsulfate (SDS), polysorbate 20 (Tween® 20), octyl phenol ethoxylate (Triton X-100), octylphenoxypolyethoxyethano (IGEPAL®-CA 630), nonyl phenoxypolyethoxylethanol (NP-40), or cetrimonium bromide (CTAB); or (e) additives selected from one or more of glycine, betaine, glycine-betaine, 7-deaza-2′-deoxyguanosine, dimethyl sulfoxide (DMSO), polyethylene glycols 400–1,000,000, glycerol, magnesium, tetramethyl ammonium chloride (TMAC), tetraethylammonium chloride (TEAC), triethylarsine hydrochloride, ethylene carbonate, dextran sulfate, or bovine serum albumin (BSA).

In another ~~embodiment~~example, the heating of the immobilized probe/target complex comprises an incubation time ranging from about 5 minutes to about 1 hour. In one ~~aspect~~example, the heating of the immobilized probe/target complex comprises an incubation temperature ranging from about

50 °C to about 70 °C. In another ~~aspect~~example, the heating of the immobilized probe/target complex comprises an incubation time of about 20 minutes and an incubation temperature of about 55 °C. In another ~~aspect~~example, the washing of the immobilized probe/target complex comprises two distinct wash buffers, wherein each wash buffer independently comprises one or more of: (a) salts selected from one or more of monovalent, divalent, sodium, ammonium, cesium, or manganese, chlorides, citrates, sulfates, perchlorates, isothiocyanates, or guanidinium; (b) chelating agents selected from one or more of ethylenediaminetetraacetic acid (EDTA), ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA), or (1,2-bis(o-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid (BAPTA); (c) buffering agents selected from one or more of tris(hydroxymethyl)aminomethane (Tris); 2-(bis(2-hydroxyethyl)amino)acetic acid) (Bicine); (*N*-[tris(hydroxymethyl)methyl]glycine (Tricine); [tris(hydroxymethyl)methylamino] propanesulfonic acid (TAPS); 3-[*N*-tris(hydroxymethyl)methylamino]-2-hydroxypropanesulfonic acid (TAPSO); 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES); 2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid (TES); 3-(N-morpholino)propanesulfonic acid (MOPS); piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES); 2-(*N*-morpholino)ethanesulfonic acid (MES); or dimethylarsenic acid (Cacodylate); (d) detergents selected from one or more of sodium dodecylsulfate (SDS), polysorbate 20 (Tween® 20), octyl phenol ethoxylate (Triton X-100), octylphenoxypolyethoxyethano (IGEPAL®-CA 630), nonyl phenoxypolyethoxylethanol (NP-40), or cetrimonium bromide (CTAB); or (e) additives selected from one or more of glycine, betaine, glycine-betaine, 7-deaza-2′-deoxyguanosine, dimethyl sulfoxide (DMSO), polyethylene glycols 400–1,000,000, glycerol, magnesium, tetramethyl ammonium chloride (TMAC), tetraethylammonium chloride (TEAC), triethylamine hydrochloride, ethylene carbonate, dextran sulfate, or bovine serum albumin (BSA).

In another ~~embodiment~~example, the washing of the immobilized probe/target complex comprises an incubation time of about 5 minutes to about 15 minutes and an incubation temperature of about 60 °C. In one ~~aspect~~example, the method improves specificity and enhances yield as compared to a conventional method. In another ~~aspect~~example, the conventional method does not comprise a step of heating prior to washing. In another ~~aspect~~example, the method improves on-target percentage as compared to a conventional method. In another ~~aspect~~example, the conventional method does not comprise a step of heating prior to washing. In another ~~aspect~~example, the method reduces handling time and throughput time as compared to a conventional method. In another ~~aspect~~example, the conventional method does not comprise a step of heating prior to washing. In another ~~aspect~~example, the method is suitable for automation and reduces complexity as compared to a conventional method. In another ~~aspect~~example, the conventional method does not comprise a step of heating prior to washing.

Another ~~embodiment~~ aspect described herein is a plurality of target nucleic acid sequences isolated using the methods described herein.

Another ~~embodiment~~ aspect described herein is a means for isolating a plurality of target nucleic acid sequences comprising any of the means, methods, steps, or compositions described herein.

Another ~~embodiment~~ aspect described herein is the use of the methods described herein to isolate a plurality of target nucleic acid sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

FIG. 1 shows a schematic of the factors affecting hybridization of single-stranded DNA with a complementary partner (e.g., a primer or probe). In comlex systems, kinetics are important. Achieving on-target duplexes by hybridization is typically slow. Using the melting step after capturing targets simplifies the workflow while increasing specificity, enhancing yields, and reducing size bias. The system is tunable by increasing the melting step temperature to increase GC-rich sequences.

FIG. 2 shows a flowchart of a standard workflow for preparation of a library for sequencing using hybridization capture.

FIG. 3 shows a schematic of a workflow for preparation of a library for sequencing using a standard hybridization capture workflow and describes the parameters at each step.

FIG. 4 shows a flowchart of the simplified “melt’ hybridization capture workflow that simplifies the steps, reduces hands-on time and turnaround time, and increases specificity, enhances yields, and reduces size bias.

FIG. 5 shows a diagram of the simplified “melt’ hybridization capture workflow with exemplary times and temperatures at each step.

FIG. 6 shows a schematic of a workflow for preparation of a library for sequencing using the simplified “melt’ hybridization capture workflow and describes exemplary parameters at each step.

FIG. 7A and 7B show a comparison of a standard hybridization capture method (xGEN SOP) or the simplified “melt’ hybridization (Short Hyb) capture method using 24-hour, 30 minute, or 15 min hybridization incubations at 65 °C. FIG. 7A shows on-target percentage and FIG 7B shows flanked on-target percentage. The targeting panel used contained 18,815 targeting probes and was hybridized with 250 ng capture input NA12878 gDNA library. The Melt-Simple Bind at