Patrik Scholler*
Steffen Heber
Jörg D. Hoheisel

Molecular-Genetic Genome Analysis Group, Deutsches Krebsforschungszentrum, Heidelberg, Germany

Optimization and automation of fluorescence-based DNA hybridization for high-throughput clone mapping

Large-scale hybridization-based genome mapping projects, such as the production of sequence-ready physical clone maps, call for robust and cheap DNA labeling techniques that are amenable to automation. We routinely use a high-throughput protocol based on fluorescence detection. DNA probes are labeled *via* polymerase chain reaction (PCR) amplification with primers that are digoxigenin-modified at their 5′ ends. Alternatively, digoxigenin-labeled dUTP is incorporated in a random hexamer priming reaction. Hybridization takes place in small volumes by sandwiching the probe between filters and plastic sheets. A fluorescence signal is produced by the activity of alkaline phosphatase attached to an anti-digoxigenin antibody upon the addition of AttoPhos™ substrate. Signals are directly detected with a charge-coupled device (CCD) camera and scored by an image data analysis system. DNA filters can be reused at least 40 times without loss of data quality. Significant advantages compared to radioactive techniques are the reduced health risk, enabling highly parallel processing; the production of spot signals uniform in size and intensity, which is essential for efficient image analysis; and a cost reduction of about 70%.

1 Introduction

Hybridization techniques based on DNA filter arrays have proved to be a powerful tool for the physical mapping of genomic DNA, the reconstruction of a genomic or chromosomal region from smaller clone fragments, by placing them relative to each other based on their hybridization behavior with various DNA probes [1, 2]. The genome of fission yeast *Schizosaccharomyces pombe*, for example, was mapped this way in yeast artificial chromosome (YAC), bacteriophage P1 and cosmid clones [3, 4]. The enormous potential of the methodology consequently led to its application in a strategy for low-redundancy sequencing by ordering shotgun template clones: the high resolution maps generated serve as a source for a directed selection of overlapping DNA templates used in the actual sequencing reactions. After demonstrating the efficiency of this process in pilot studies on chromosome XII of budding yeast *Saccharomyces cerevisiae* [5–7], the technique was optimized and applied to sequencing projects in the megabase range of human and microbial DNA (Scholler *et al.*, manuscripts in preparation). For this purpose, radioactive labeling was replaced by fluorescence-based detection of hybridization signals and the protocols were adapted to permit high-throughput data production (Fig. 1).

Fluorescence-based DNA labeling is a well-established methodology in assays in which signal intensities or signal-to-background ratios are sufficiently high, such as fluorescence *in situ* hybridization (FISH; signal detection with fluorescence microscope; [8]) or automated DNA sequencing (excitation by high-energy monochromatic light source; [9]). Many different fluorochromes were developed to meet various specifications [10]. For filter-based DNA hybridization, with many signal spots to be detected simultaneously, several aspects are crucial: after stimulation with a scattering, surface-covering light source, the fluorescence signal intensity must be high and stable enough for unequivocal detection, with a CCD camera or an analogous detection device; the signal-to-noise ratio (commonly used nylon filters contain fluorescent components themselves) as well as the homogeneity of the signal intensities must allow clear spot definition and reliable evaluation; and last, but no less important, filters must be reusable in many subsequent experiments. The fluorophor AttoPhos™ [11], with its long lifetime and a relatively large Stokes shift of 140 nm, in combination with alkaline phosphatase for enzymatic signal amplification, is an ideal candidate [12]. Here, such fluorescence-based protocols are described, adapted, and optimized for routine application in physical clone mapping.

2 Materials and methods

2.1 Filter and probe preparation

Filter arrays of 80 clones/cm^2 were produced from size-selected shotgun template libraries with a commercial robotic device (BioRobotics, Cambridge, UK). DNA was attached to the filters for repeated use in hybridization experiments as described in detail elsewhere [4, 5]. Non-radioactive DNA probes were either generated from restriction fragments or from the shotgun-clone inserts by random hexamer priming [13] or a standard PCR amplification. To both reaction types, digoxigenin-11-dUTP (Boehringer Mannheim, Mannheim, Germany) was added at a concentration of 10 μM. Alternatively,

Correspondence: Dr. J. D. Hoheisel, Molecular-Genetic Genome Analysis Group, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany (Tel: +49-6221-42-4680; Fax: +49-6221-42-4682; E-mail: j.hoheisel@dkfz-heidelberg.de)

Keywords: Fluorescence labeling / Genome mapping / Hybridization / Image analysis

* Current address: LION Bioscience AG, Im Neuenheimer Feld 517, D-69120 Heidelberg, Germany

0173-0835/98/0404-0504 $17.50+.50/0

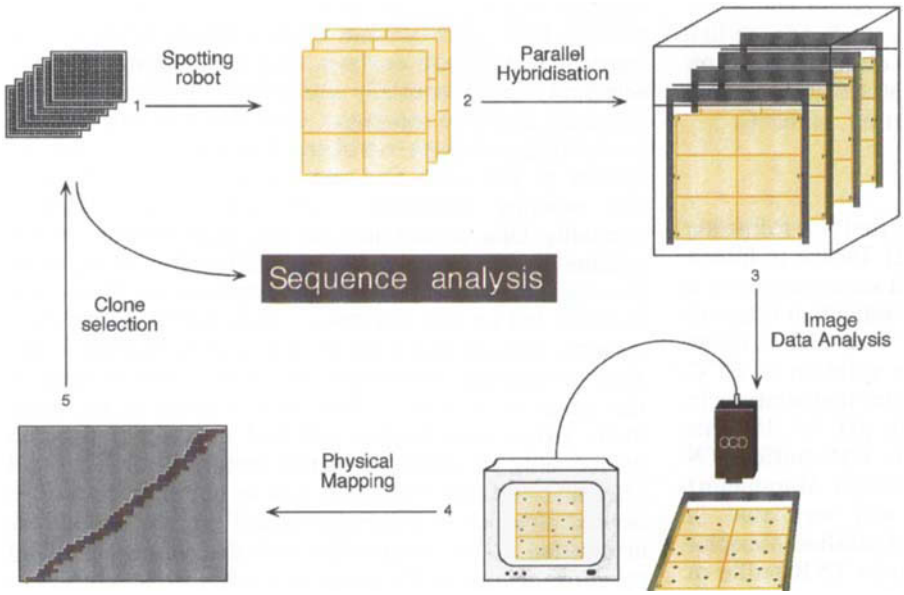

Figure 1. Procedures involved in high-resolution physical genome mapping: (1) Conventionally made, size-selected shotgun libraries are spotted to nylon filters of 22 × 22 cm using a robotic device. (2) Filters are repeatedly used in hybridization experiments. (3) Fluorescence signals are detected with a CCD camera system connected to an appropriate image analysis software. (4) Signal spots are transformed to clone names and automatically used for map generation. (5) By algorithms that are individually adaptable to the sequencing performance (*e.g.*, average read-length), clones are selected from the map for ordered DNA sequence analysis.

cloning-site flanking primers modified with digoxigenin at their 5′ position were used in PCR, delivering results of equivalent quality. Clone-inserts were PCR-amplified in sealed 384 well plates in an Autogene II cycler (Grant, Cambridge, UK). Wells were filled with 25 μL of 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM of each nucleotide triphosphate, 0.01 mM digoxigenin-11-dUTP, 1 μM of each primer, and 1 unit of *Taq* DNA polymerase using a multichannel plate filling device (*e.g.*, Genetics, Christchurch, UK) prior to inoculation with bacterial or M13 phage suspension transferred from a culture plate by a 384 pin replicator (Genetics). After an initial incubation at 95°C for 2 min, 32 PCR cycles were performed by alternating a 2 min period of annealing and primer extension at 68°C and 30 s of denaturation at 95°C. Post-treatment was at 72°C for 10 min. PCR products were checked and simultaneously purified by brief electrophoresis in 1% w/v agarose gels so that the fragments all traveled nearly identical distances. Using the microplate array diagonal gel electrophoresis (MADGE) system [14], material was transferred to the gel in the microtiter dish format. DNA bands were either excised with a home-made, fitted tool that simultaneously dissects all gel areas known to contain the PCR fragments, or selectively in ethidium bromide-stained gels under long-wave UV light, and were directly used for hybridization. Pooling of many PCR products as mixed probes was possible without loss of data quality.

2.2 Hybridization

For probe release and denaturation, agarose blocks containing the digoxigenin-labeled DNA were heated to 100°C for 6 min in the presence of 50 μL water and immediately mixed with 2 mL of 0.5 M sodium phosphate, pH 7.2, 7% w/w SDS, 1 mM EDTA, pH 8, 100 μg/mL sonified salmon sperm DNA or tRNA, prewarmed to 65°C. Filters were attached at their four corners to specifically designed metal frames (Fig. 2) for easy handling during the washing steps. Hybridization was performed either in compact hybridization boxes that consist of several narrow chambers for individual hybridization experiments (Engineering and Design Plastics, Cambridge, UK) or by uniformly spreading 4 mL of

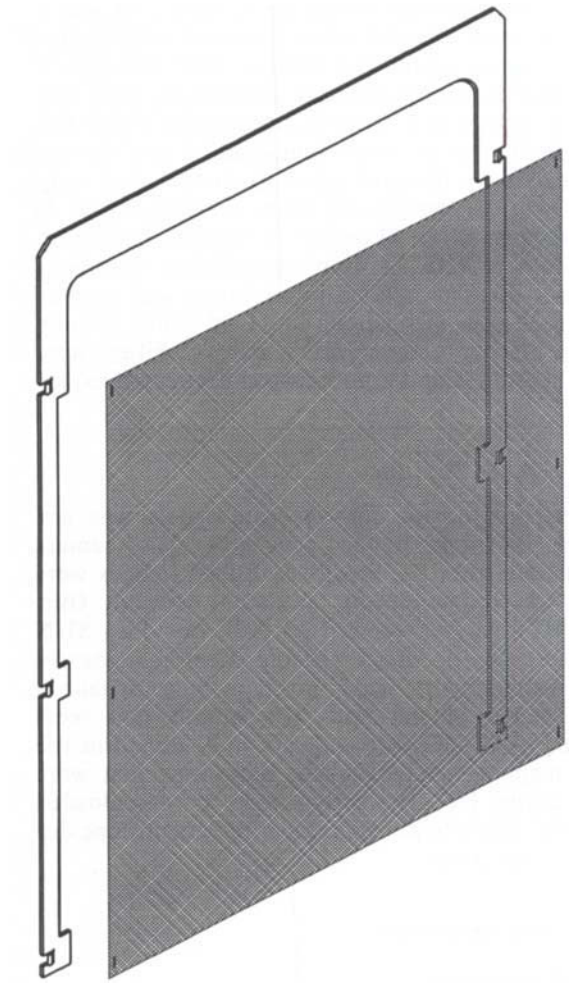

Figure 2. Scheme of the metal frame to which the filters were attached by small perforations for simple handling during all processing steps.

probe between a filter surface of 22 × 22 cm and a plastic sheet of slightly larger dimensions. Sandwich-like stacks of 10–15 such arrangements – directly placed on top of each other – were wrapped in plastic foil to avoid evaporation during overnight incubation at 65°C.

2.3 Fluorescence detection

Washing and incubation steps were carried out following a protocol modified by Maier *et al.* [12]. Twelve to fifteen filters hybridized with different probes were washed in a plastic box. After a short rinse of the individual filters in 40 mM sodium phosphate, pH 7.2, 0.1% w/v SDS, filters were washed in 500 mL of the same solution at 65°C, followed by 5 min incubation at room temperature in 500 mL TN buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl) and another 5 min in 500 mL TBN-buffer (TN plus 1% w/v blocking reagent, Boehringer Mannheim). Subsequently, the entire filter set was incubated in 300 mL of 0.05 U/mL anti-digoxigenin alkaline phosphatase complex (Boehringer Mannheim) in TNB buffer at room temperature for about 30 min. The antibody solution was reused 5 times over a period of about 4 h. Prior to incubation with the fluorescence substrate, filter sets were submerged first in 500 mL of TN buffer, subsequently twice in 500 mL of 100 mM Tris-HCl, pH 9, 100 mM NaCl, 1 mM $MgCl_2$, and finally in 500 mL of 100 mM diethanolaminechloride, pH 9, and 1 mM $MgCl_2$. Filters were briefly blotted onto Whatman 3MM paper to remove excess liquid and sprayed with AttoPhos™ solution (Boehringer Mannheim). Incubation, while protected from drying by a cover of transparent plastic foil, was for 30–90 min at room temperature, and fluorescence signals were detected upon irradiation with long-wave UV light. To remove the probe material, up to 30 filters were washed in 1 L of 0.5 M sodium phosphate, pH 7.2, 5% w/v SDS, at room temperature overnight. After a brief rinse in distilled water, probe was stripped off by two 30 min incubations in 1 L of 5 mM sodium phosphate, pH 7.2, 0.1% w/v SDS, at 95°C. Filters were stored in hybridization buffer between sequential experiments.

2.4 Data analysis

Signal detection of the fluorescence signals was performed by a charge-coupled device (CCD) camera system directly from the processed filters. Images were imported into commercially available software (BioImage HDG Analyzer, version 2.1 [15]) run on a SUN Sparc-5 workstation, which reliably identifies positive spots by matching the actual image with an idealized reference image of the grid. The output files were reformatted by a self-made software tool, and form the basis for mapping analysis using algorithms that were described earlier [16]. The software can be downloaded freely from http://www.mpimg-berlin-dahlem.mpg.de/~andy/welcome.html.

3 Results and discussion

3.1 Parallel processing

For establishing high-throughput mapping, a system was set up for parallel handling of hybridization experiments (Fig. 1). At the current status, the crucial steps – clone picking, PCR amplification, DNA spotting, hybridization, signal detection and analysis, map generation and clone selection – are essentially automated, mostly using commercially available machines. Only the non-rate limiting, connecting processes – moving dishes or filters from one station to the next, and the addition of hybridization and washing solutions – are still being performed manually. One critical step for the establishment of the scheme was the substitution of radioactive isotopes by fluorescence labeling. For hybridizations on filters, our decision fell on the AttoPhos™ system [12] for practical reasons: lack of exposure to radioactivity facilitates parallel processing; differences in probe concentration in the range of at least 15-fold were leveled by the enzymatic signal amplification and had no visible effect on data quality in terms of signal uniformity in size and intensity; and cost reduction was by approximately 70%. Several steps of an existing protocol [12] were modified, or omitted, to be compatible with the parallel handling of, currently, up to 60 filters at a time, while simultaneously achieving reproducible data quality at the lowest possible cost. Concerning the latter, the concentration of the anti-digoxigenin alkaline phosphatase conjugate, which is a significant cost factor, was reduced by 75%, for example, in addition to the fact that it was reused 4–5 times. Filters lasted for at least 40 subsequent experiments, mostly longer. These data are based upon hybridization of more than 2000 probes.

3.2 Fluorescence detection

We have had negative experience with laminating filters on one side [17] for better transportation: they tended to deform after probe stripping at 95°C. A specifically designed flat metal frame (Fig. 2) that holds a filter at its four corners proved to be useful. Held by these frames, the filters could be placed easily in boxes that are subdivided into chambers for individual hybridizations or washing steps. In most cases, however, hybridization was done in small volumes of solution that were spread between the filter and a sheet of plastic. Up to 15 such sandwiched arrangements were stacked for hybridization. Subsequently, they were washed simultaneously in a large volume of buffer, after an initial rinse of individual filters in slots of a hybridization box.

Signal detection of the fluorescence signals was performed by a CCD camera system directly from the processed filters. Illumination of the area was done by two UV lamps (length, 30 cm; wavelength, 365 nm; power, 15 W) placed about 30 cm above either edge of the filter at a slight, inward angle. This arrangement, chosen for its technical simplicity as compared to laser-based scanning devices for example, nevertheless produced excellent results (Fig. 3). For coordinate identification, images were imported to a software package (BioImage HDG Analyser, version 2.1), which identifies positive spots by matching the actual image with an idealized reference image of the grid [15]. Where low-quality hybridization data had to be analyzed, which was avoided if possible, a semi-manual scoring method was used instead. On a computer screen, a reference grid was pulled to the correct size to fit the relevant data image. Positives, judged

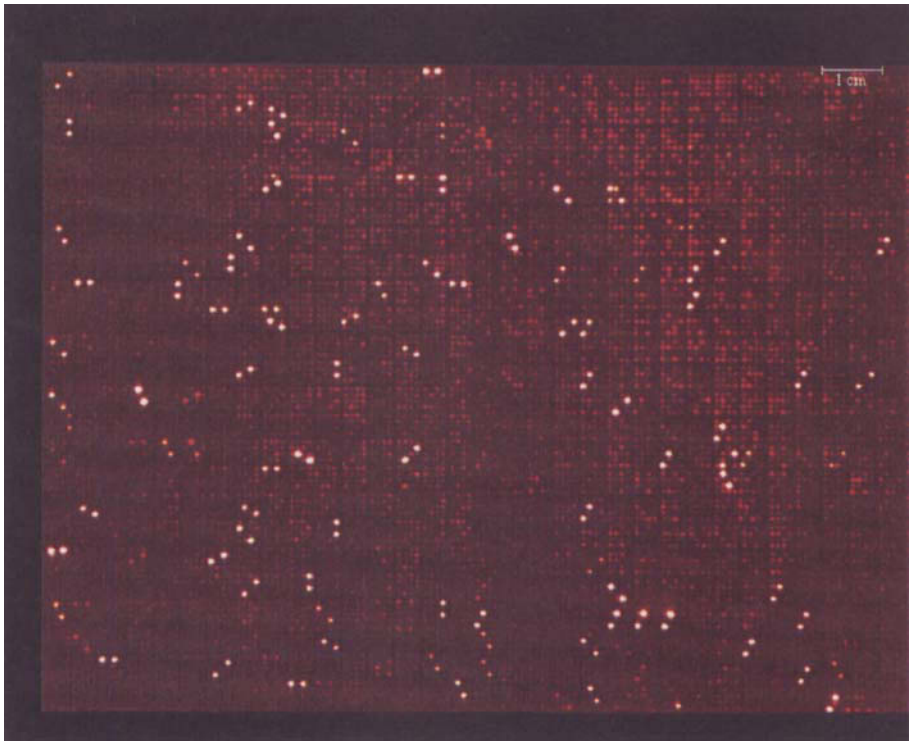

Figure 3. Typical result of a hybridization to a filter clone array with a fluorescently labeled probe of pooled clone inserts made by PCR. The spots were arranged in blocks of 4 × 4 at an average density of 80 clones per cm^2. Within a block, each DNA was present twice. Apart from serving as an internal control of positive hybridization, the orientation of the related spots within the box revealed the DNA's identity.

by the human eye, were scored by a mouse click at the respective spots of the reference grid.

3.3 Assessment of data quality

Data quality was assessed manually and the files were then used for mapping, using an established program package [16] that calculates probe distance values and computes a minimal distance spanning all probes. For improved analysis, two extra features were added: listing the clones not positive in any hybridization, which by definition should be located in yet unmapped regions and thus lend themselves as probes in subsequent hybridizations, and automatic selection of clones for the definition of a minimal clone tiling path. The distance of probes is calculated by a maximum likelihood approach whereupon the shortest spanning clone coverage is sought through a two-dimensional graph, with the additional requirement that every second clone had been used as a probe. Thereby, gaps in clone coverage are avoided.

3.4 Further developments

The current set-up is an initial phase toward a more complete, automated system for clone mapping by hybridization experiments. The modular organization was chosen for practical reasons, such as the fact that commercial devices could be used, and also because of financial considerations. The remaining manual manipulation is not rate-limiting in the first instance. Nevertheless, developments are being pursued toward further automation of these handling steps, which could well be performed by commercially available robotics. At present, a filter that is worth 0.5 Mbp of shotgun clones, for example, can be handled simultaneously in 60 copies. By hybridizing pools of six probes per filter, some 5 to 6 rounds of hybridization are necessary to produce the data for a complete template map of such a 0.5 Mbp region.

4 Concluding remarks

Clone mapping is considered by some a receding art because of developments such as sequencing strategies based on clone selection by end sequencing [18] or even total shotgun approaches [19] when in fact it offers a way out of the unnecessary cost factor of high sequence redundancy, which, due to improved chemistry, is no longer a prerequisite to accurate sequencing. Rather than decreasing the mapping effort, we embarked on the opposite. The directed approach taken reduces the redundancy of finished sequence to an overall value of well below 4-fold, about half that of a standard shotgun activity, and adds extra value by a significantly lowered complexity in sequence assembly and a simplified finishing phase. Essential to the technology, however, are protocols for achieving the required throughput. With the current system, both mapping cost and time could be reduced to about 10% of the corresponding sequencing effort. Although already a substantial improvement, this difference by a factor of (only) ten clearly indicates that prospective gains are still to be made in mapping procedures, which, after all, deal with kilobase-sized fragments instead of base pairs.

For microorganisms, with their short intergenic stretches, a minimal clone tiling path made from a template map also represents a complete and perfectly normalized gene inventory. RNA profiling experiments [20] on such

a substrate could render the sequencing of an entire microbial genome unnecessary. Only genes and related genomic regions that exhibit an interesting response would be chosen for analysis. By this selective approach, only a few tens of thousands of base pairs of high potential would be sequenced rather than millions of bases of unqualified importance.

Received July 31, 1997

5 References

[1] Lehrach, H., Drmanac, R., Hoheisel, J. D., Larin, Z., Lennon, G., Monaco, A. P., Nizetic, D., Zehetner, G., Poustka, A., in: Davies, K. E., Tilghman, S. (Eds.), *Genome Analysis Vol. 1: Genetic and Physical Mapping,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1990, pp. 39–81.

[2] Hoheisel, J. D., *Trends Genet.* 1994, *10*, 79–83.

[3] Maier, E., Hoheisel, J. D., McCarthy, L., Mott, R., Grigoriev, A. V., Monaco, A. P., Larin, Z., Lehrach, H., *Nature Genet.* 1992, *1*, 273–277.

[4] Hoheisel, J. D., Maier, E., Mott, R., McCarthy, L., Grigoriev, A. V., Schalkwyk, L. C., Nizetic, D., Francis, F., Lehrach, H., *Cell* 1993, *73*, 109–120.

[5] Scholler, P., Karger, A. E., Meier-Ewert, S., Lehrach, H., Delius, H., Hoheisel, J. D., *Nucleic Acids Res.* 1995, *23*, 3842–3849.

[6] Scholler, P., Benes, V., Voss, H., Ansorge, W., Hoheisel, J. D., *DNA Sequence* 1996, *6*, 257–262.

[7] Johnston, M., Hillier, L., Riles, L. *et al., Nature* 1997, *387*, 87–90.

[8] Joos, S., Fink, T. M., Rätsch, A., Lichter, P., *J. Biotechnol.* 1994, *35*, 121–153.

[9] Ansorge, W., Voss, H., Wiemann, S., Schwager, C., Sproat, B., Zimmermann, J., Erfle, H., Hewitt, N., Rupp, T., *Electrophoresis* 1992, *13*, 616–619.

[10] Paddock, S. W., *Curr. Biol.* 1997, *7*, 182–185.

[11] Cano, R. J., Torres, M. J., Klem, R. E., Palomares, J. C., *BioTechniques* 1992, *12*, 264–268.

[12] Maier, E., Crollius, H. R., Lehrach, H., *Nucleic Acids Res.* 1994, *22*, 3423–3424.

[13] Feinberg, A. P., Vogelstein, B., *Anal. Biochem.* 1983, *132*, 6–13.

[14] Day, I. N. M., Humphries, S. E., *Anal. Biochem.* 1994, *222*, 389–395.

[15] Granjeaud, S., Nguyen, C., Rocha, D., Luton, R., Jordan, B. R., *Genet. Anal.* 1996, *12*, 151–162.

[16] Mott, R., Grigoriev, A. V., Maier, E., Hoheisel, J. D., Lehrach, H., *Nucleic Acids Res.* 1993, *21*, 1965–1974.

[17] Bancroft, D. R., O'Brien, J. K., Guerasimova, A., Lehrach, H., *Nucleic Acids Res.* 1997, *25*, 410–416.

[18] Smith, M. W., Holmsen, A. L., Wei, Y. H., Peterson, M., Evans, G. A., *Nature Genet.* 1994, *7*, 40–47.

[19] Fleischmann, R. D., Adams, M. D., White, O. *et al., Science* 1995, *269*, 496–512.

[20] Schena, M., Shalon, D., Davis, R. W., Brown, P. O., *Science* 1995, *270*, 467–470.