doi: 10.1038/nature06458

Silicon Nanowires as Highly Efficient Thermoelectric Materials—Supplementary Information (Experiment)

Akram I. Boukai#, Yuri Bunimovich#, Jamil Tahir-Kheli, Jen-Kan Yu,

William A. Goddard III, and James R. Heath*

Device Fabrication and Measurement Setup

The devices used in this report were fabricated from intrinsic silicon-on-insulator (SOI) substrates (Ibis Technology) with a buried oxide thickness of 150nm. The top Si(100) epilayer began with a thickness of 400nm and was thinned through repeated thermal oxidations and wet chemical etching steps until the desired Si layer thickness (20nm or 35nm) was obtained. AFM analysis revealed a smooth surface with an rms roughness of ~3.6 Å. Next, the Si epilayer was doped using a boron containing spin-on dopant (Filmtronics Boron-A) (21). After spin coating, the substrates were baked (250 ºC; 5 min) and annealed under flowing N_2 using a rapid thermal processor at temperatures corresponding to the desired doping concentration by allowing the boron to diffuse into the silicon. Four-point probe conductivity measurements determined the p-type doping concentration. We have previously reported that the doping profile through the thickness of the epilayer drops rapidly from the top surface (Figure S1).

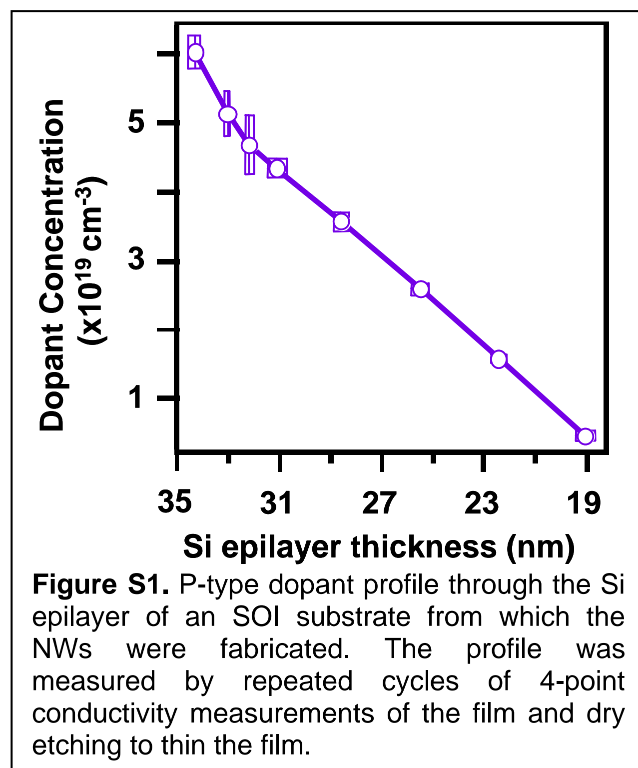

Figure S1. P-type dopant profile through the Si epilayer of an SOI substrate from which the NWs were fabricated. The profile was measured by repeated cycles of 4-point conductivity measurements of the film and dry etching to thin the film.

After the diffusion doping process, the substrates were used to fabricate Si NWs using the SNAP process (20). The depth of the NWs was nominally 20nm (the SOI substrate thickness). However, the “electrical” depth, the distance at which the doping concentration decreased by an order of magnitude or more, was 10-15 nm (Fig. S1). SEM and AFM were used to determine the

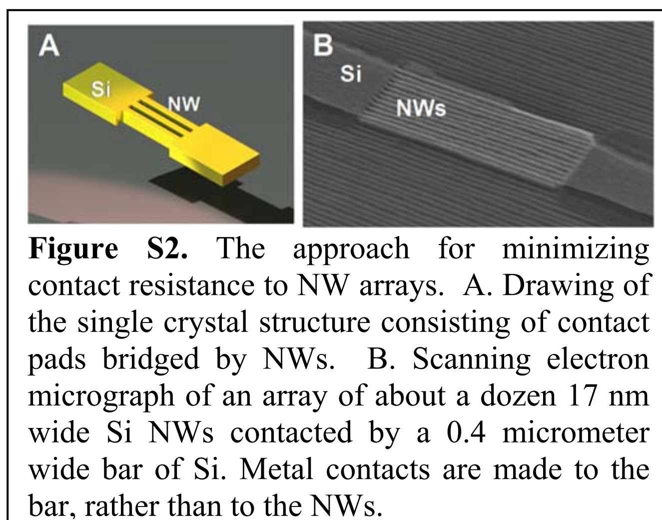

Figure S2. The approach for minimizing contact resistance to NW arrays. A. Drawing of the single crystal structure consisting of contact pads bridged by NWs. B. Scanning electron micrograph of an array of about a dozen 17 nm wide Si NWs contacted by a 0.4 micrometer wide bar of Si. Metal contacts are made to the bar, rather than to the NWs.

dimensions of our Si NWs. Devices contained from 10 to 400 NWs connected in parallel. For all measurements on the 10 nm wide wires other than thermal conductivity, monolithic contacts were used in order to minimize the contact resistance (Fig S2.). The use of monolithic contacts means that the NWs constitute a bridge between two micrometer-sized pads, with the NWs and pads all part of the same single crystal silicon structure. This method effectively removes contact resistance.

Measurement of the three thermoelectric parameters, **S**, σ, and κ was accomplished using a device architecture similar to those of other published reports (12, 16, 31, 36) (Fig. S3). Heaters and electrodes were fabricated using e-beam lithography (EBL) and e-beam evaporation of metal. The heaters and electrodes were both composed of titanium/platinum and were 20nm/100nm thick and 20nm/160nm thick respectively. Prior to metallization of the electrodes, the Si NWs were dipped in a buffered HF solution for 3 seconds to remove any native oxide. After electrode metallization, the device was annealed in forming gas inside the RTP (475°C; 5 min). The electrodes were then contacted by large gold pads (10nm/220nm Ti/Au) that were defined by photolithography.

We used EBL and several etching steps to suspend the Si NWs. Polymethyl methacrylate (PMMA), an EBL resist material, is spun coated onto the surface of a chip containing a fully fabricated device. EBL is used to define an opening in PMMA around the active area of the device. A small rectangular island of PMMA remains on top of the NWs, electrodes and heaters to protect them during subsequent etching steps. Next, we used a CF_4 plasma etch, 20 minutes at 40W and 20 mTorr, to remove the SiO_2 around the active region. A 1.5 minute XeF_2 etch removes the exposed silicon and undercuts the silicon handle layer underneath the active region. During this step, the NWs are protected by the thick rectangular island of PMMA on top and the 150nm thick oxide underneath. The PMMA is then removed using an O_2 plasma (5 min; 40W;

20 mTorr). At the end, the Si NWs remain anchored on top of a thin rectangular island of 150nm thick SiO_2.

Removal of the oxide island would cause the NWs to collapse, and so, as described below, measurements of κ were carried out as a differential measurement – i.e. a measurement of the NWs + oxide island followed by a measurement of the oxide island. Since this is a measurement of the heat flow only through the NWs, the oxide acts as a parasitic path for this heat flow that must be experimentally quantified.

Electrodes contacting the Si NWs were initially shorted together to minimize device damage through static charge. When the measurement begins, the shorts were disconnected. The Si substrate is mounted onto a thermally conductive chip carrier with the gold pads wire-bonded to that same chip carrier. The chip carrier is then mounted onto the cold finger of a vacuum sealed liquid nitrogen cryostat (Janis Research VPF-475), which is evacuated to $\leq 10^{-6}$ Torr for the measurements. A temperature controller (Lakeshore Model 331) allows global control of the cryostat temperature from 350K to 77K with a precision of .01K.

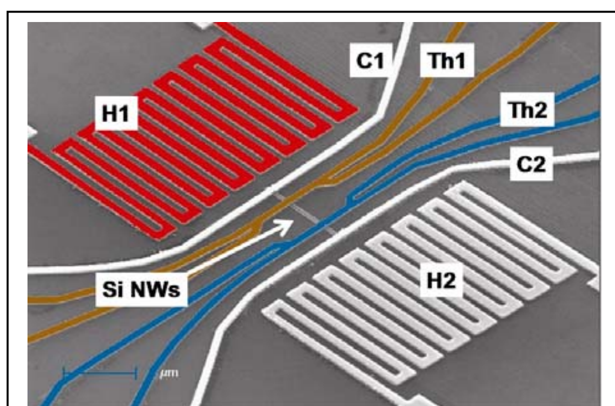

Figure S3. False color SEM image of a (non-suspended) thermoelectric device platform used in this study. The heaters (H1 and H2) are shown, one in red to indicate that Joule heating is done using just one of the heaters. C1, C2, Th1 and Th2 comprise the 4-point electrical contacts to the Si NWs, with C1 and C2 utilized as the current source and drain for those electrical conductivity measurements. Th1 and Th2 also serve as resistive thermometers (note that these two electrodes are themselves 4-point devices).

Thermopower: Measurement and Data

Determining the thermopower, $\frac{\Delta V}{\Delta T}$, requires measurements of the thermoelectric voltage ΔV and the temperature difference ΔT along the Si NWs. A switching matrix (Keithley 707A) allows reconfigurable electrical connections from device to instrumentation for the various thermoelectric measurements.

We first create a temperature gradient along the Si NWs using a DC current through one heater (Joule heating) (Fig. S3). We then measure the resultant thermoelectric voltage, ΔV (TEV). A Keithley 2182A nano-voltmeter measures the TEV between the inner electrodes (Th1 and Th2). In order to verify that the measured voltage is truly a result of the temperature gradient, the measurement is repeated using the second heater situated on the opposite side of the NW circuit. This should yield a voltage with similar magnitude but opposite polarity. A second check is to reverse the current polarity in the heater – this should not alter the magnitude or the polarity of the thermoelectric voltage. Fig. S4A verifies that the measured voltage is indeed due to the temperature gradient.

Two separate measurements yield ΔT by using the inner electrodes as resistive thermometers. A resistive thermometer responds to temperature changes with a proportionate change in resistance. In the first measurement, the resistance of the Ti/Pt electrodes is continuously recorded by a four-point probe measurement as a function of heater power with one heater on. As the power dissipated in the heater slowly ramps stepwise, four lock-in amplifiers (Stanford Research Systems SR-830) simultaneously monitor the resistance of both thermometers. The first pair of lock-in units measures the AC current and voltage at 913 Hz for one thermometer. Another pair of lock-in units performs the same functions for the other thermometer at 921 Hz. We obtain the resistances, R, of each thermometer as a function of heater power, Q (Fig. S4B).

In the second measurement, we ramp the cryostat temperature stepwise by 1K increments while simultaneously measuring the thermometer's resistance. We obtain a linear relationship between the change in resistance and the cryostat temperature (Fig. S4C). This relates our thermometer's resistance change to a temperature difference, ΔT, through the relation:

$$\Delta T = W \bullet \left(\frac{dR_{Left}/dQ}{dR_{Left}/dT} - \frac{dR_{Right}/dQ}{dR_{Right}/dT} \right)$$

where R_{Left} and R_{Right} are the resistances of the left and right thermometers respectively. The slope of a linear fit to the TEV vs ΔT data provides the value of the thermopower. This process is repeated to calculate the thermopower from 300K down to 100K.

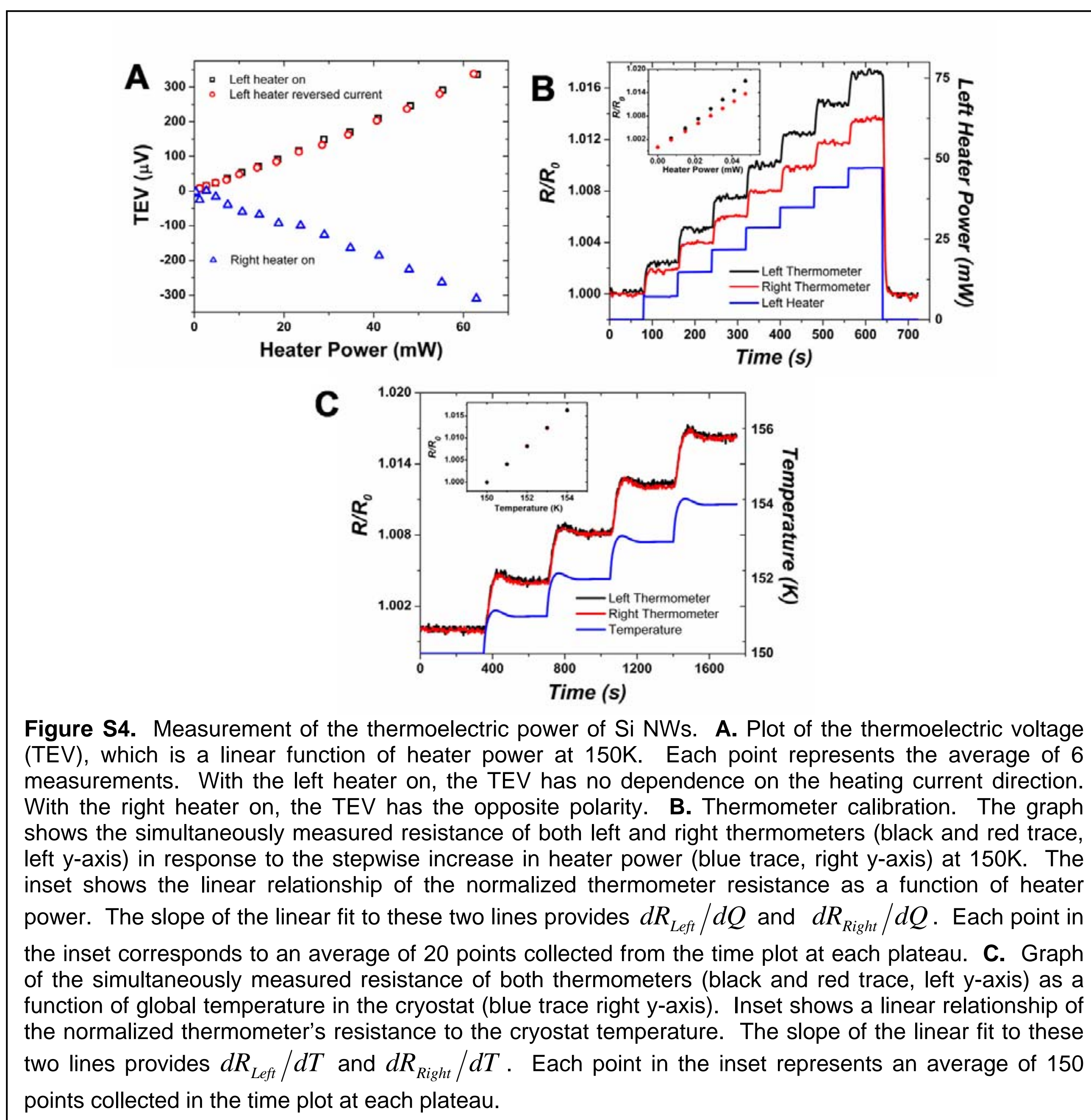

Figure S4. Measurement of the thermoelectric power of Si NWs. **A.** Plot of the thermoelectric voltage (TEV), which is a linear function of heater power at 150K. Each point represents the average of 6 measurements. With the left heater on, the TEV has no dependence on the heating current direction. With the right heater on, the TEV has the opposite polarity. **B.** Thermometer calibration. The graph shows the simultaneously measured resistance of both left and right thermometers (black and red trace, left y-axis) in response to the stepwise increase in heater power (blue trace, right y-axis) at 150K. The inset shows the linear relationship of the normalized thermometer resistance as a function of heater power. The slope of the linear fit to these two lines provides dR_{Left}/dQ and dR_{Right}/dQ. Each point in the inset corresponds to an average of 20 points collected from the time plot at each plateau. **C.** Graph of the simultaneously measured resistance of both thermometers (black and red trace, left y-axis) as a function of global temperature in the cryostat (blue trace right y-axis). Inset shows a linear relationship of the normalized thermometer's resistance to the cryostat temperature. The slope of the linear fit to these two lines provides dR_{Left}/dT and dR_{Right}/dT. Each point in the inset represents an average of 150 points collected in the time plot at each plateau.

Thermopower: Supplementary Data

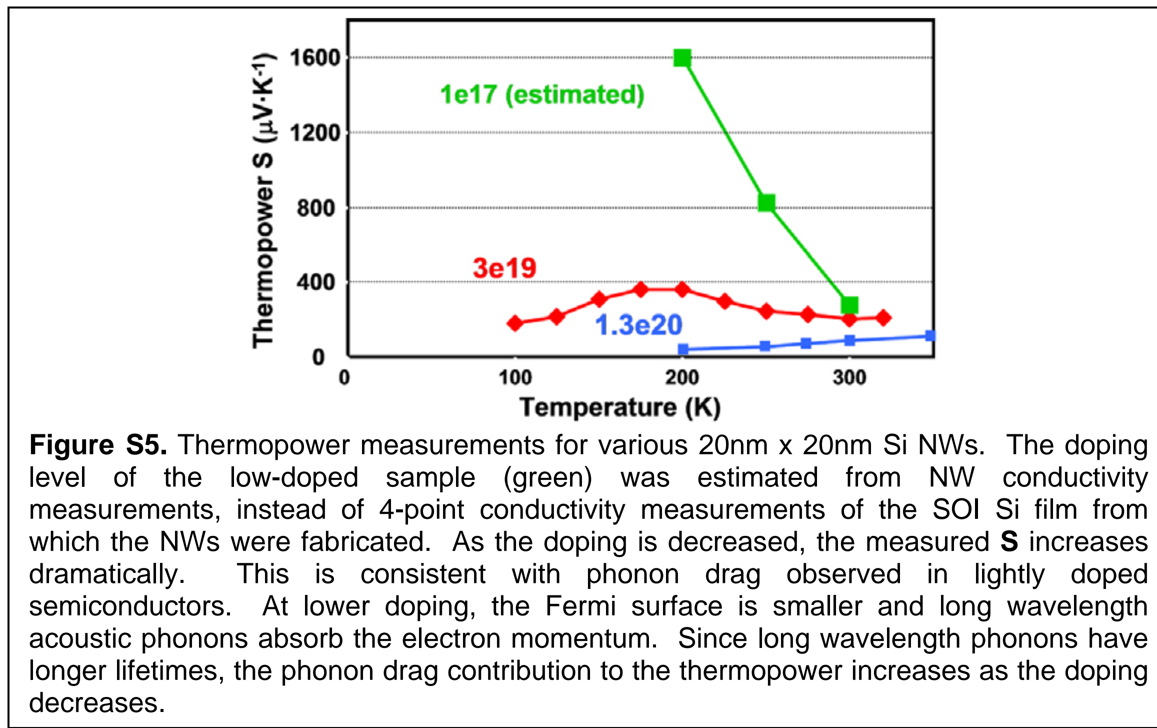

Figure S5. Thermopower measurements for various 20nm x 20nm Si NWs. The doping level of the low-doped sample (green) was estimated from NW conductivity measurements, instead of 4-point conductivity measurements of the SOI Si film from which the NWs were fabricated. As the doping is decreased, the measured **S** increases dramatically. This is consistent with phonon drag observed in lightly doped semiconductors. At lower doping, the Fermi surface is smaller and long wavelength acoustic phonons absorb the electron momentum. Since long wavelength phonons have longer lifetimes, the phonon drag contribution to the thermopower increases as the doping decreases.

Thermal Conductivity: Measurement and Data

Techniques utilized to measure the thermal conductivity, κ, were developed by Majumdar's group (17,18). The general expression for the thermal conductance of any system is given by the expression: **Q**=**K**ΔT, where **Q** is the heat transferred across the NWs, ΔT is the temperature difference and **K** is the thermal conductance (**K** is related to κ through geometrical factors). Measurement of κ is not as straightforward since parasitic pathways for the heat flow, **Q**, must be accounted for. Therefore, we determine the parasitic heat flow through the beams that support the suspended structure. A direct measurement of the power dissipated in the heater, $\mathbf{Q_{Heater}}$, is possible by using the 4-point probe electrodes attached to the heater's windings since the power dissipated in the heater is equal to the product of its current and voltage drop. A lock-in measurement similar to that described for measurement of the thermopower allows us to measure ΔT. Also, the thermometers are in intimate contact with the ends of the NWs, eliminating any contributions from contact thermal resistances. The resistance of the thermometers is typically two orders of magnitude smaller than the resistance of the NWs. Taking into account the thermal conductance of the beams, $\mathbf{K_b}$, the thermal conductance of the sample, $\mathbf{K_s}$, is given by (17):

$$K_s = K_b \frac{T_C}{\Delta T} \text{............eq. 1}$$

where T_C is the temperature increase at the

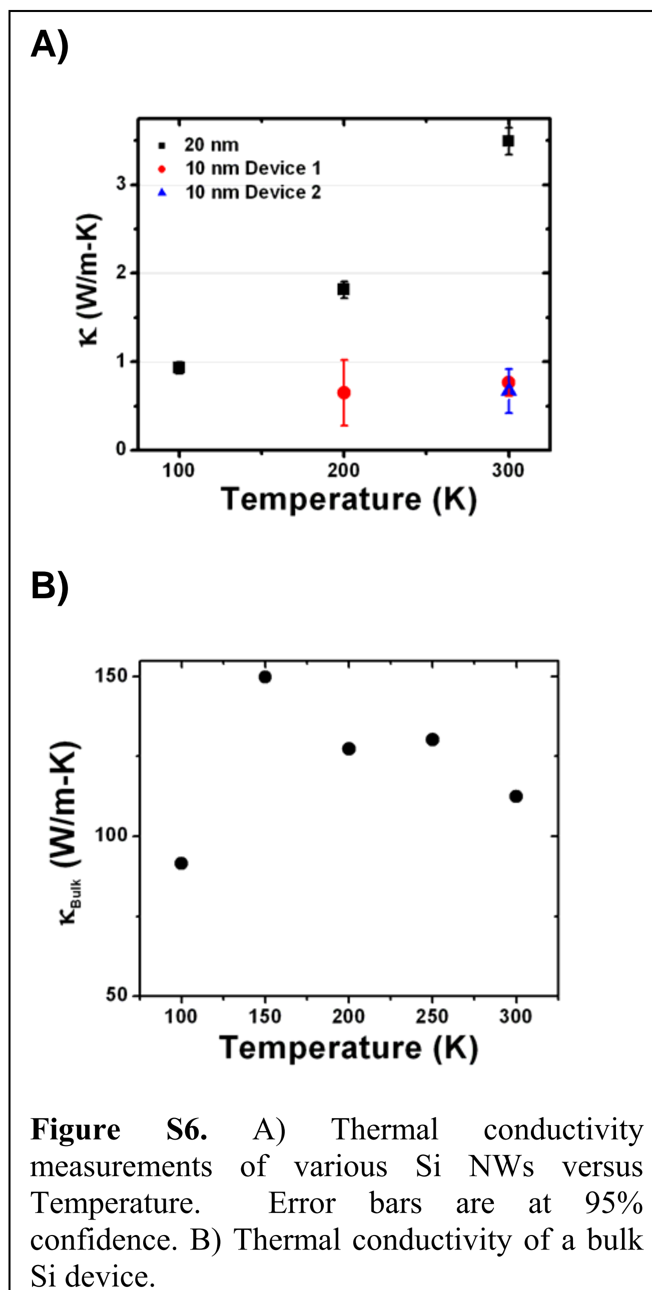

Figure S6. A) Thermal conductivity measurements of various Si NWs versus Temperature. Error bars are at 95% confidence. B) Thermal conductivity of a bulk Si device.

cold junction and ΔT is the temperature difference between the hot and cold junctions. The thermal conductance of the beams is given by:

$$K_b = \frac{Q}{T_H + T_C} \text{............eq. 2}$$

where Q is the power dissipated in the heater and T_H is the temperature increase of the hot junction.

Our suspended structure is a combination of the SiO_2 island that supported the NWs and electrodes (Fig 1), and the Si NWs. To estimate the contribution of the SiO_2 to the measured $\mathbf{K_{total}}$, we selectively remove the NWs with a XeF_2 etch (32) (Figure 2) and then measure $\mathbf{K_{SiO2}}$. The thermometers were separately calibrated during each thermal conductance measurement – both before and after the Si NWs were removed. $\mathbf{K_{SiO2}}$ is subtracted from $\mathbf{K_{total}}$, giving $\mathbf{K_{NWs}}$. Finally, $\mathbf{K_{NWs}}$ is converted to the $\boldsymbol{\kappa}$ from dimensions and number of NWs (Fig. S6). The $\boldsymbol{\kappa_{bulkSi}}$ ~150 W $m^{-1}K^{-1}$ at 300K. $\boldsymbol{\kappa}$ of thin films of Si has been found to be reduced by 50-70% over the bulk, depending upon the thickness, doping, and roughness of the film (35). Our measured value of ~100 W $m^{-1}K^{-1}$ at 300K is somewhat higher than other reported values. This may arise from factors such as the smoothness of our epilayers or the doping profile through the thickness of our films. Our measured values are consistent with published reports from Majumdar's and Wang's groups (18). Also, the change in the thermal conductance of the metallic beams before and after XeF_2 is much less than 1%. This indicates that the difference between the two values of the thermal conductance arises solely from removal of the NWs.

Electrical Conductivity: Measurement and Data

The electrical conductivity of our NWs was measured using a four-point probe technique that removes any contribution from the contacts. A Keithley 2400 provides a current source through the outer electrodes (Figure 2) and the nano-voltmeter measures the voltage drop between the two inner electrodes. The conductance value is calculated from the slope of a linear fit of the I-V data. The conductance is then converted to an electrical conductivity using the dimensions & numbers of the NWs (Fig. S7).

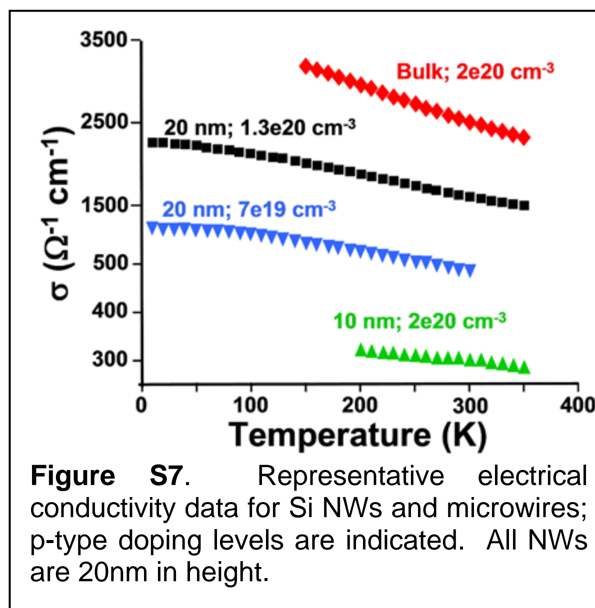

Figure S7. Representative electrical conductivity data for Si NWs and microwires; p-type doping levels are indicated. All NWs are 20nm in height.

The measured thermopower includes the contribution from the oxide substrates. We show that the contribution from the oxide thermopower S_{oxide} is negligible because the oxide resistance is much larger than the NWs resistance.

Since our measured S_{total} = 200 μV/K and S_{oxide} = 400 μV/K at 300K (37), the oxide contribution to the measured thermopower is negligible when the resistance of the oxide is much larger than the resistance of the NWs. This can be seen by the equation below for S_{total}

$$S_{Total} = \frac{G_{oxide}S_{oxide} + G_{NW}S_{NW}}{G_{oxide} + G_{NW}}$$

where G_{oxide} and G_{NW} are the conductances of the oxide and NWs respectively. The resistance of our NWs is ~ 10^4 Ω at 300K.

To determine the resistance of our oxide substrate, we performed a current-voltage measurement on the substrate. Since we did not obtain a linear fit to the data (due to the tiny currents <100pA at ~ 1 V), we estimated the order of magnitude of the resistance by dividing each individual current-voltage point. This gave us a resistance order of magnitude of ~ 10^{10} Ω Since the substrate resistance is ~ 10^6 times larger than the NWs resistance and the orders of magnitude of

their respective thermopowers is the same, the total thermopower, S_{total}, is completely dominated by the NW contribution.

Error Analysis- Thermal Conductivity

A measurement of the thermal conductivity of our Si NWs requires knowing the power dissipated in the heater and the temperature difference across the NWs. In addition, we must determine the parasitic heat flow through the beams that support the suspended structure. Heat also flows through the thin oxide film underneath the NWs. The measured thermal conductivity value is then a contribution from the NWs + oxide. To eliminate the contribution from the oxide, we remove the NWs using a highly selective XeF_2 etch and determine the thermal conductivity of the oxide in a second measurement. The difference between the two measured values is the thermal conductivity of the NWs.

The thermal conductance of the sample, $\mathbf{K_s}$, is given by (17):

$$K_s = K_b \frac{T_C}{\Delta T} \text{............Equation 1}$$

where K_b is the thermal conductance of the beams, T_C is the temperature increase at the cold junction after the heater is turned on and ΔT is the temperature difference between the hot and cold junctions. The thermal conductance of the beams, K_b, is given by:

$$K_b = \frac{Q}{T_H + T_C} \text{............Equation 2}$$

where Q is the power dissipated in the heater and T_H is the temperature increase of the hot junction after the heater is turned on. Using error propagation, the error of K_s, δK_s, is give by:

$$\delta K_s = K_s \sqrt{\frac{\delta K_b^2}{K_b^2} + \frac{\delta T_C^2}{T_C^2} + \frac{[\delta(T_H - T_C)]^2}{(T_H - T_C)^2}} \text{............Equation 3}$$

Further simplifying this expression, we substitute for the error of K_b and ΔT:

$$\frac{\delta K_b^2}{K_b^2} = \frac{\delta Q^2}{Q^2} + \frac{\delta T_H^2 + \delta T_C^2}{(T_H + T_C)^2}, \text{............Equation 4}$$

$$[\delta(T_H - T_C)]^2 = \delta T_H^2 + \delta T_C^2 \text{............Equation 5}$$

$$\delta K_s = K_s \sqrt{\frac{\delta Q^2}{Q^2} + \frac{\delta T_C^2}{T_C^2} + 2\frac{(\delta T_H^2 + \delta T_C^2)(T_H^2 + T_C^2)}{(T_H^2 - T_C^2)^2}} \text{............Equation 6}$$

The error in Q can be ignored because its uncertainty, $\frac{\delta Q}{Q}$, is 1 part in 1000000 or .001%. The errors in the temperature measurement for both hot and cold junctions are therefore the dominant terms in equation 6. Determining the error of the temperature measurement requires using error propagation on the formula used to find the temperature. The thermopower measurement section describes how we measured the temperature of each junction. The temperatures of the hot and cold junction are given by:

$$T_H = Q \bullet \left(\frac{dR_H / dQ}{dR_H / dT} \right) \qquad T_C = Q \bullet \left(\frac{dR_C / dQ}{dR_C / dT} \right)$$

where R_H and R_C are the resistances of the hot and cold thermometers. The term, $\left(\frac{dR / dQ}{dR / dT} \right)$ is called the conversion factor and is calculated for each thermometer. The terms, dR/dQ and dR/dT, are the slopes of the best linear fits to the data points in the resistance vs. heater power and resistance vs. temperature plots respectively (see thermopower measurement section).

For example, each data point on the resistance vs. heater power plot is determined by selecting a point from each plateau in the resistance vs. time plot for the heater power calibration. Figure 1 shows how this is accomplished. Since there are 55 plateaus and 100 useable points/plateau after temperature equilibration, we generate a series of 100 lines, with each line defined by 55 points. The slopes of all 100 lines are calculated and averaged giving dR/dQ. The error of the average slope is calculated by finding the standard deviation of all 100 slopes and dividing by $\sqrt{100}$. This value is:

$$\delta\left(dR / dQ \right) = \frac{\sigma}{\sqrt{100}}$$

where σ is the standard deviation of the 100 slope values. Similarly, $\delta\left(dR / dT \right)$ is calculated by generating 800 linear fits, each with 40 data points (since there are 800 useable points/plateau after equilibration in the resistance vs. time plot and 40 plateaus for the temperature calibration).

The value of dR/dT is found by taking the average of all 800 slopes and the error is

$$\delta\left(dR/dT\right) = \frac{\sigma}{\sqrt{800}}.$$

We now find the error for our temperature measurement by first calculating the error of the conversion factor. The error in the conversion factor is:

$$\delta\left(\frac{dR/dQ}{dR/dT}\right) = \left(\frac{dR/dQ}{dR/dT}\right)\sqrt{\left[\frac{\delta\left(dR/dQ\right)}{dR/dQ}\right]^2 + \left[\frac{\delta\left(dR/dT\right)}{dR/dT}\right]^2}$$

Therefore, the error in the measured temperature for both thermometers is:

$$\delta T = T\sqrt{\left[\delta\left(\frac{dR/dQ}{dR/dT}\right) \Big/ \frac{dR/dQ}{dR/dT}\right]^2}$$

where we have ignored the error in **Q** because it is small. Based on the data collected, our uncertainty in the temperature is .01% at 300K. This sets the uncertainty in $\mathbf{K_s}$ to .068%. We obtain two values for $\mathbf{K_s}$, before etching the NWs and after etching, and then subtract these to value to obtain $\mathbf{K_{NW}}$. The error in $\mathbf{K_{NW}}$ is:

$$\delta K_{NW} = \sqrt{\delta K^2_{s\ Before\ Etch} + \delta K^2_{s\ After\ Etch}}$$

We have a total error of .47 nW/K (or 9.7%) at 300K for the thermal conductance of our 10nm NWs. Finally, we determine the thermal conductivity, $\kappa = K_s\, L/A$ where L and A are the length and cross-sectional area of our NWs respectively. The uncertainty of the thermal conductivity is:

$$\delta\kappa = \kappa\sqrt{\left(\delta K_s/K_s\right)^2 + \left(\delta L/L\right)^2 + \left(\delta A/A\right)^2} \quad \text{...........equation 7}$$

The length, L, is measured by scanning electron microscopy (SEM) and its error is ~5nm, while A is measured by AFM and its error is .1nm. Therefore the dominant term in the right-hand side of equation 7 is the uncertainty of $\mathbf{K_s}$. The percent uncertainty of our thermal conductivity value at 95% confidence for our 10nm NWs is 2(δκ/κ) = 19.5% at 300K.

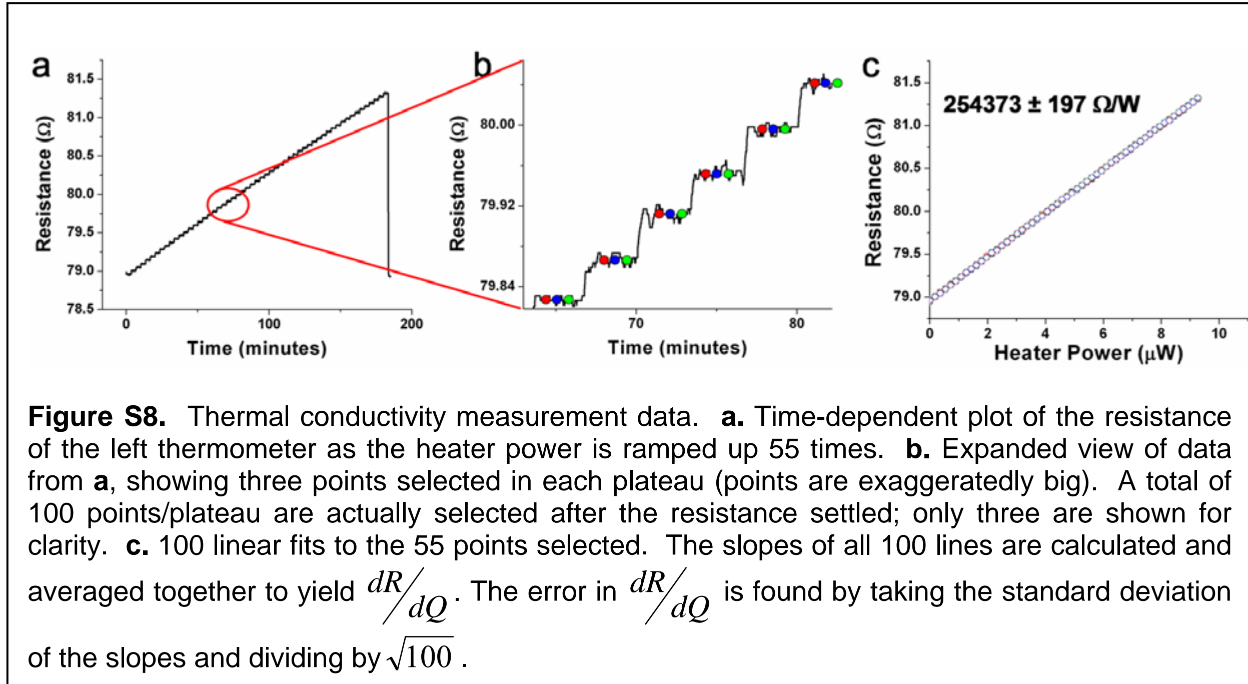

Figure S8. Thermal conductivity measurement data. **a.** Time-dependent plot of the resistance of the left thermometer as the heater power is ramped up 55 times. **b.** Expanded view of data from **a**, showing three points selected in each plateau (points are exaggeratedly big). A total of 100 points/plateau are actually selected after the resistance settled; only three are shown for clarity. **c.** 100 linear fits to the 55 points selected. The slopes of all 100 lines are calculated and averaged together to yield dR/dQ. The error in dR/dQ is found by taking the standard deviation of the slopes and dividing by $\sqrt{100}$.

Error Analysis- Thermopower and Electrical Conductivity

The measured **S** (= $\frac{\Delta V}{\Delta T}$) is the slope of a best fit line to ΔV as a function of ΔT. The error of each ΔV point was found by calculating the standard deviation of each of the six voltage measurements taken for that point (Fig. S3A). The relative error was found to be ~ 1% whereas for lower doped NW's, the relative error was ~ 5%.

Next, the error of the temperature difference, δT, was determined using the same method as stated above for the thermal conductivity error analysis. Since we need to determine the temperature difference, ΔT, the error is:

$$\delta T \approx \sqrt{\left[\frac{dR_H/dQ}{dR_H/dT}\cdot\sqrt{\left(\frac{\delta\left(dR_H/dQ\right)}{dR_H/dQ}\right)^2+\left(\frac{\delta\left(dR_{Ht}/dT\right)}{dR_H/dT}\right)^2}\right]^2+\left[\frac{dR_C/dQ}{dR_C/dT}\cdot\sqrt{\left(\frac{\delta\left(dR_C/dQ\right)}{dR_C/dQ}\right)^2+\left(\frac{\delta\left(dR_C/dT\right)}{dR_C/dT}\right)^2}\right]^2}$$

where we have neglected the error in Q since it is less than 1%. The relative error for ΔT is found to be 2-3% for most devices measured throughout the full temperature range (300K to 100K).

Finally, the uncertainty in **S** is calculated from the standard error of the slope of the best fit lines to the ΔV vs. ΔT plot at the 95% confidence level. We find that our relative error for **S** is between 1and 6%.

The error of the electrical conductivity is less than 1% and the current-voltage traces are highly linear with $R^2 \sim .999$.

Error Analysis- ZT

The uncertainty in ZT is calculated using the following formula:

$$\delta ZT = ZT\sqrt{\left[4\left(\frac{\delta S}{S}\right)^2 + \left(\frac{\delta\kappa}{\kappa}\right)^2 + \left(\frac{\delta\sigma}{\sigma}\right)^2\right]}.$$

The factor of four appears in the above equation because $ZT \propto S^2$. The relative error in σ can be ignored since it is less than 1%. The uncertainty in ZT at the 95% confidence level, $2\times\delta ZT$, is shown in Figure 3 of the main paper for two different devices.

Supplementary Information (Theory)

1D versus 3D electronic structure The first question to consider in understanding the unexpectedly large Seebeck effect observed for the relatively lightly doped NWs is whether the electronic structure in the 10-20 nm wires is 1D or 3D above 100K. This can be estimated from the magnitude of the zero-point energy due to the finite cross-section of the wire. The effective masses for the two bands at the Valence Band Maximum of Si are 0.49 and 0.16 electron masses. Hole doping will occur predominantly in the heavier band. The zero-point energy is

$$\varepsilon = \frac{\hbar^2 k^2}{2m^*}, \quad k = \frac{\pi}{d},$$

where d is the wire diameter. For m* = 0.49, ε = 89K for d = 10 nm and ε = 22K for d = 20 nm. Thus for temperatures above 100K thermal smearing is larger than the zero-point energies, suggesting that the electronic properties should be 3D. Furthermore, the thermopower data in Figs. 2 (green and pink data points) and 3 (blue data points) shows that the 10 and 20nm wires and bulk samples with similar impurity doping levels have an almost identical magnitude and linear temperature dependence. Since these 10 nm, 20 nm, and bulk wires have similar thermopowers, the different zero-point energies are not leading to a large change in **S** and we may conclude the electronic structure in all the wires remains 3D for temperatures > 100K.

Is there a Boron impurity band?

For Boron doping levels within the range of 10^{19}-10^{20} cm^{-3}, the electron gas becomes degenerate and it is possible that a B impurity band forms containing the charge carriers. By comparing an estimate of the thermopower due to holes in the Si valence bands of the NW with measured values, we conclude that this is unlikely.

If the holes are primarily in the Si valence bands, then the holes are predominantly in the heavier valence band with mass 0.49. For doping n, the Fermi momentum satisfies

$$n = \frac{k_f^3}{3\pi^2}.$$

For n = 3 x 10^{19} cm^{-3}, k_f = 0.1 $Å^{-1}$ leading to Fermi energy E_f = 0.072 eV = 833K. From the Mott formula for the electronic contribution **S**:

$$S_e(T) = \frac{\pi^2 k^2 T}{3e}\left(\frac{\partial \ln \sigma(\varepsilon)}{\partial \varepsilon}\right) \approx (283\,\mu V / K)(kT / E_f),$$

we estimate that $\mathbf{S_e}$ = 102 μV/K at T = 300K. We assumed that the derivative of the logarithm of the conductivity is the reciprocal of the energy scale over which $\boldsymbol{\sigma}(\varepsilon)$ changes ~ $1/E_f$ in computing $\mathbf{S_e}$. This estimate for $\mathbf{S_e}$ is compatible with $\mathbf{S_e}$ = 130-170 μV/K from experiments for bulk samples at 300K for wires that only have a linear T electronic contribution to **S**. A Boron impurity band would likely lead to much heavier masses and hence to $\mathbf{S_e}$ larger than observed.

Derivation of the results from elasticity theory for the phonon lifetime

For doping n = 3 x 10^{19} cm^{-3}, the longitudinal acoustic (LA) phonons that can receive momentum from the electrons are characterized by wavelengths of > 31 Å, corresponding to > 5.5 unit cells of Si. Classical elasticity theory models the system by an elastic continuum. Indeed, recent molecular dynamics simulations show classical elasticity is a good approximation for these wavelengths (27). Since the wavelengths of the phonons involved in phonon drag are larger than several unit cells, classical elasticity theory is expected to give a good approximation to their dynamics. When a sound wave propagates through an elastic medium it generates local temperature gradients from the local volume changes in the wave. Thermal conduction acts to remove these gradients by heat flow with a corresponding increase in entropy. This irreversible heat flow leads to damping of the sound wave (26).

Consider a volume element δxδyδz = δx δA with a temperature gradient along the x-direction. The temperature at x is T and T + δT at x + δx. Over a time interval δt an amount of heat δQ

$$\delta Q = (+\kappa \nabla T)\delta A\,\delta t$$

flows to T. **κ** is the thermal conductivity. The change in entropy (*S*) is

$$\delta S = \frac{\delta Q}{T} - \frac{\delta Q}{T + \delta T}.$$

The irreversible energy lost due to the increasing entropy is

$$T\delta S = \frac{\delta T}{T}\delta Q,$$

leading to the rate of energy lost by thermal conduction

$$\frac{dE}{dt} = -\frac{\delta T}{T}\frac{\delta Q}{\delta t} = -\frac{\delta T}{T}\kappa\frac{\delta T}{\delta x}\delta y\delta z = -\frac{\kappa}{T}\left(\frac{\delta T}{\delta x}\right)^2 dV,$$

$$\frac{dE}{dT} = -\kappa\int\frac{1}{T}(\nabla T)^2 dV.$$

A 3D longitudinal wave given by $u_x = u_0\cos(kx-\omega t)$ where $\omega=c_L k$ and c_L is the longitudinal speed of sound leads to local temperature gradients due to local volume changes as the wave propagates. The volume deformations of the longitudinal wave are separated by one-half the wavelength λ. For low ω, the separation between the high and low temperature regions is too large to be equalized during the period of the wave. Thus for low ω, the volume changes are adiabatic (33). At high frequencies, the wavelength is small and thermal gradients are rapidly equalized leading to isothermal volume changes for large ω (33).

A 3D longitudinal acoustic (LA) phonon with momentum (q_x, q_y, q_z) that makes an angle Θ with the long axis of the 1D wire (z-direction) is a linear combination of a longitudinal phonon of momentum q_z with amplitude cosΘ and a transverse acoustic (TA) phonon with momentum q_z with amplitude sinΘ. Thus electron-phonon scattering around the Fermi surface leads to a 1D LA mode with probability $\cos^2\Theta$ and a TA mode with probability $\sin^2\Theta$. TA modes have temperature gradients perpendicular to the wire axis. LA modes have temperature gradients along the wire axis.

The largest momentum (shortest wavelength) phonon modes participating in phonon drag by being excited due to the holes scattering with phonons are 3D longitudinal acoustic with wavevector, $k_{LA} = 2k_f = 0.2$ $Å^{-1}$ for an impurity doping level of 3×10^{19} cm^{-3}. The wavelength is $\lambda_{sh} = 2\pi/k = 31$ Å. The speed of sound in the (100) direction for Si is $c_L = 8.43 \times 10^5$ cm/s, leading to $\omega_{LA} = c_L k = 1.7 \times 10^{13}$ sec^{-1}. Phonon drag thermopower is an average over all the phonon modes that are scattered by the electrons at the Fermi surface. Thus ω_{LA} is an upper bound on the phonon frequencies participating in phonon drag.

The crossover frequency from adiabatic volume deformations to isothermal deformations for LA modes occurs at $\omega_0(LA) = c_L^2/\chi$ where χ is the thermal diffusivity. $\chi = 0.8$ cm^2/s for bulk Si at 300K leading to $\omega_0(LA) = 8.9 \times 10^{11}$ sec^{-1} (28). 3D LA modes with frequency ω_{LA} are isothermal.

For TA modes, the crossover frequency is determined by the width of the wire, d, and the thermal diffusivity, $\omega_0(TA) = \pi^2\chi/d^2$. In this case, frequencies *below* $\omega_0(TA)$ are isothermal and frequencies *above* $\omega_0(TA)$ are adiabatic (28, 29). For d = 20nm, $\omega_0(TA) = 2 \times 10^{12}\ s^{-1}$ using the bulk value of the thermal diffusivity.

For the NW, the thermal diffusivity is ~100 times smaller than bulk since the diffusivity is proportional to the thermal conductivity. Thus the LA crossover frequency is ~100 times larger than the bulk $\omega_0(LA)$ leading to adiabatic damping for LA modes. The TA mode crossover frequency decreases by a factor of ~ 100. TA modes are also adiabatic in the NW at 300K.

The damping rate for an LA sound wave of momentum k is (33, 34)

$$\frac{1}{\tau_{LA}(k)} = \frac{1}{\tau_0} \cdot \frac{(\omega_k \tau_0)^2}{1+(\omega_k \tau_0)^2} \cdot \varepsilon,$$

where the terms in the expression are defined as

$$\omega_k = c_L k, \qquad \tau_0 = \frac{\chi}{c_L^2}, \quad \chi = \frac{\kappa}{\rho C_p}, \quad \varepsilon = \frac{v_{th}^2}{c_L^2}, \quad v_{th}^2 = \left[\frac{\alpha^2 K_{ad}^2 T}{\rho C_p}\right].$$

ρ is the density, χ is the thermal diffusivity, $\boldsymbol{\kappa}$ is the thermal conductivity, α is the volume thermal expansion coefficient ($\alpha = 3\alpha_{linear}$ where α_{linear} is the linear expansion coefficient) , K_{ad} is the adiabatic bulk modulus, c_L is the speed of sound, and C_p is the constant pressure specific heat per volume. For $\omega_k \gg 1/\tau_0$, the damping saturates to ε/τ_0. This occurs for 3D LA modes relevant to phonon drag.

The ratio of the 1D and 3D LA scattering times for a mode with momentum k is

$$\frac{\tau_{3D}(k)}{\tau_{1D}(k)} = (kc_L)^2 \left(\frac{\chi}{c_L^2}\right)^2 \left[\frac{(1-\nu)}{(1+\nu)(1-2\nu)}\right]\left[\frac{\kappa(1D)}{\kappa(3D)}\right],$$

where ν is the Poisson ratio ($\nu = 0.29$ for Si). For the largest momentum $k = 2k_f$, the 3D LA mode lifetime is 3.3 times larger than the 1D LA mode. We take the ratio of the thermal conductivities to be 1/100. For the shortest wavelength LA modes, phonon drag is not enhanced in the wire. For long wavelength LA modes (small k), the lifetime of the 1D mode rapidly

increases relative to the bulk due to the k^2 term and phonon drag is enhanced. In addition, the 3D scattering rate for long wavelength LA modes becomes adiabatic rather than isothermal leading to a further increase in the 1D LA mode lifetime relative to 3D. In this regime, the lifetime is increased by the ratio of the 3D to 1D thermal conductivities or ~ 100.

The scattering rate of the 1D TA modes is smaller than the corresponding LA modes in bulk. The scattering rate of the 1D TA mode is (33)

$$\frac{1}{\tau_{TA}(k)} = \omega_0(TA) \cdot \frac{(\omega_k / \omega_0(TA))^2}{1 + (\omega_k / \omega_0(TA))^2} \cdot \Delta, \qquad \Delta = \left[\frac{\alpha_{linear}^2 K_{ad} T}{C_p} \right]$$

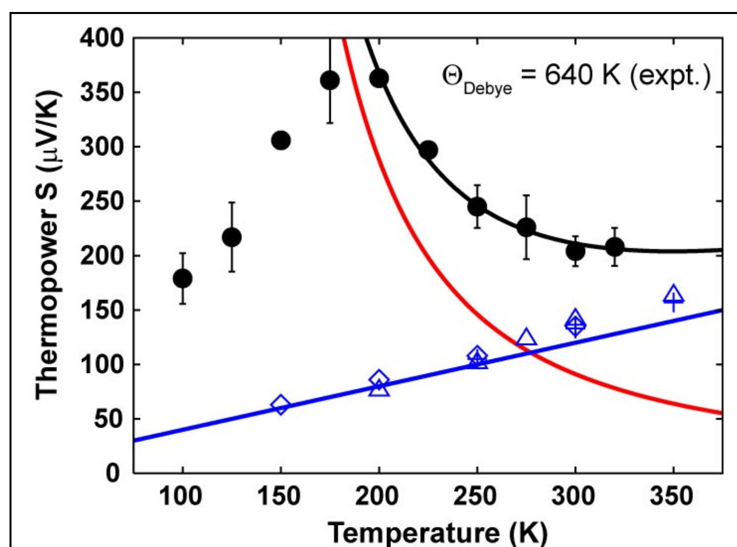

Figure S8. A refit of Fig. 4 using the experimentally determined Θ_{Debye} = 640K. The points are the experimental data also presented in Fig. 4. The black, red and blue lines represent the fitted curves of **S**, $\mathbf{S_{ph}}$, and $\mathbf{S_e}$, respectively, utilizing the expression in the text. We find, a = 0.40 $\mu V/K^2$ and b = 12.2 μV/K. The fitted maximum and rms error are 8.9 μV/K and 2.1 μV/K respectively. The experimental error bars at 150, 200, and 225K are smaller than the data points. The fitted electronic coefficient value is close to our estimated value of a = 0.34 $\mu V/K^2$ assuming hole doping of Si with no boron impurity band. It is also close to the blue symbols in this figure which are for samples that only exhibit a linear T $\mathbf{S_e}$ component.

Since $\omega >> \omega_0(TA)$, the TA modes are adiabatic and the scattering rate saturates to $\Delta\omega_0$. Using experimental values K_{ad} = 9.8 x 10^{11} dyne/cm, α_{linear} = 2.6 x 10^{-6} K^{-1}, and ρ = 2.3 g/cm^3 leads to τ_{3D}/τ_{TA} ~ 1/2. The phonon drag contribution from the TA modes is increased due to the decreased thermal conductivity in the NW.

Thus we have shown that the decreased thermal conductivity leads to increased phonon drag due to TA modes and long wavelength LA modes.

Supplementary Figure (Theory)

In Figure S8 we show how well the high temperature data (>200K) in Fig. 4 is explained by:

$$S = S_e + S_{ph} = aT + b\left[\exp\left(\Theta_D / T\right) - 1\right]$$

where we set the Debye energy parameter to its experimental value of Θ_{Debye} = 640K. As shown in Figs. 4 and S7, phonon drag explains the observed monotonic increase in **S** as T decreases.

Ultimately, we must have **S**→0 as T →0 (by the 3rd law of thermodynamics) and this is observed experimentally. One factor that we have not included is that at sufficiently low temperature, the mean free path approaches the dimensions of the system leading to scattering that eventually reduces **S** toward zero. It is also possible that for the doping levels used here the energy dependence of the electron-phonon scattering reduces the momentum transfer to LA phonons leading to a further reduction in $\mathbf{S_{ph}}$ at low T.

References – Suplementary Information

31. Llaguno, M.C., Fischer, J.E., Johnson A.T. & Hone, J. Observation of thermopower oscillations in the coulomb blockade regime in a semiconducting carbon nanotube. *Nano Lett.* **4**, 45-49 (2004).

32. Williams, K.R. & Muller, R.S. Etch rates for micromachining processing. *Journal of Microelectromechanical Systems* **5**, 256-269 (1996).

33. Deresiewicz, H. Plane waves in a thermoelastic solid. *Journal of the Acoustical Society of America* **29**, 204-209 (1957).

34. Chadwick, P. & Sneddon, I.N. Plane waves in an elastic solid conducting heat. *Journal of the Mechanics and Physics of Solids* **6**, 223-230 (1958).

35. Liu, W.J. & Asheghi, M. Thermal conduction in ultrahigh pure and doped single-crystal silicon layers at high temperatures. *Journal of Applied Physics* **98**, 123523 (2005).

36. Zhou, F. *et. al.*, Determination of transport properties in chromium disilicide nanowires via combined thermoelectric and structural characterizations. *Nano Letters* **7**, 1649-1654 (2007).

37. Beynon, J. & Steele, C.B. Thermopower measurements on SiO_x thin films. Journal of Materials Science 25, 4255-4258 (1990).